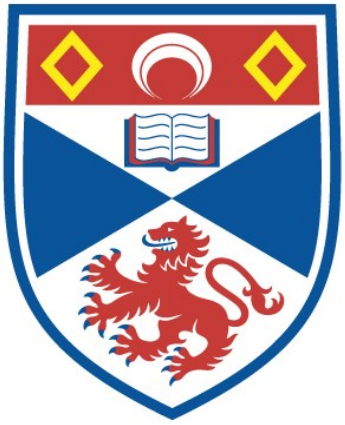
SYNTHETIC AND STRUCTURAL STUDIES OF FOOT-AND-MOUTH DISEASE VIRUS POLYPROTEIN PROCESSING

Morag May Lenman

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1996

SYNTHETIC AND STRUCTURAL STUDIES OF FOOT-AND-MOUTH DISEASE VIRUS POLYPROTEIN PROCESSING

a thesis presented by

Morag May Lenman

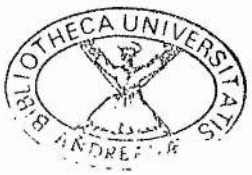
to the

UNIVERSITY OF ST. ANDREWS

in application for

THE DEGREE OF DOCTOR OF PHILOSOPHY

St. Andrews August 1995



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

DECLARATION

I, Morag Lenman, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed ... Date 21-8-95

I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No. 12 on 01-10-92 and as a candidate for the degree of Ph.D. on 01-10-93

Signed . Date 21-8-95

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the degree of Ph.D.

Signature of supervisor

Date 28 Aug 1995

Copyright

TO MY PARENTS

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor David Gani for his ideas, enthusiasm and for providing such good working facilities. I would also like to thank Dr. Mahmoud Akhtar for his invaluable help in the preparation of this thesis, Dr. Martin Ryan for helpful discussions on the biology front and Dr. Alan Aitken for his kindness and helpful discussions on the chemistry front.

I am also indebted to Dr. Ian Sadler (Edinburgh University) for running the 600 MHz NMR spectra, Dr. Nick Price (Stirling University) for recording the CD spectra and Dr. Scott Ingham (Edinburgh University) for performing the crystallographic studies.

I would like to express thanks to all my colleagues in lab 434 for providing such a pleasant working atmosphere and wish them all well in their future careers.

Above all, however, I would like to thank my parents for the unstinting generosity and support which they have shown to me, not only throughout my postgraduate studies but over the last two and a half decades. This thesis is dedicated to them.

ABSTRACT

The 2A region of the foot-and-mouth disease virus (FMDV) polyprotein is only 16 amino acids in length. During synthesis of the FMDV polyprotein, a primary proteolytic processing event occurs between the 2A and 2B regions of the polyprotein at a Gly-Pro junction. Since the 2A region is too small to be an enzyme, and there is evidence to rule out host protease involvement, it is proposed that the 2A segment represents a self-cleaving system.

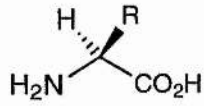
Synthetic oligopeptides containing the 2A sequence have been prepared using solid phase peptide synthesis and their three-dimensional structures in different solvents have been investigated, using NMR and CD techniques. Computer modelling studies of the 2A region have been carried out and have indicated the presence of a *cis* prolyl bond. Short peptide fragments containing the sequence of the 2A region around the scissile bond have also been prepared by solution phase peptide synthesis and their conformations examined, using NMR spectroscopy. In particular, the NPGP tetrapeptide, which is alleged to possess cleavage activity was prepared and its reactivity investigated.

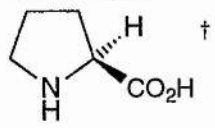
The influence of prolyl *cis/trans* isomerism on the structure of polypeptides is poorly understood and a synthetic approach has been used to prepare *cis* X-Pro peptides. The strategy involved the formation of bicyclic compounds containing a hydrazide linkage with a view to breaking the N-N bond under mild conditions to give the natural *cis* X-Pro peptide. The design, synthesis and reactivity of various novel, bicyclic *cis* X-Pro dipeptides is described. These compounds represent a new class of type VI β-turn mimetics and as it has been shown that they can be easily extended at the amino and carboxy termini for incorporation into larger peptides. X-ray crystal structures for some of the compounds and intermediates are presented. Investigations into the reductive cleavage of these compounds were carried out and some were found to display unusual reactivity with the chosen reduction method. In particular, a novel intramolecular transamidation reaction is reported and its mechanism has been investigated by the use of various substituted derivatives.

Abbreviation	Meaning
^{t}BOC	*tert*-butoxycarbonyl
CD	circular dichroism
CBZ	benzyloxycarbonyl
COSY	correlation spectroscopy
DKP	diketopiperazine
DMAP	4-dimethylaminopyridine
DMF	*N,N'*-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
FMDV	Foot-and-mouth disease virus
E. coli	Escherichia coli
Fmoc	9-fluorenylmethoxycarbonyl
h	hour
HPLC	high performance liquid chromatography
IR	infra-red spectroscopy
KAPA	potassium-3-aminopropyl amide
min	minutes
M_r	relative molecular mass
NMM	*N*-methyl morpholine
NMR	nuclear magnetic resonance
NOBA	sodium *ortho*-benzoate
nOe	nuclear Overhauser enhancement
NOESY	nOe correlation spectroscopy
PyBOP	benzotriazolyloxy-tris[pyrrolidino]-phosphonium hexafluorophosphate
PyBroP	bromo tripyrrolidino phosphonium hexafluorophosphate
RNA	ribonucleic acid

mRNA	messenger ribonucleic acid
ROESY	rotating frame nOe correlation spectroscopy
RPHPLC	reversed-phase high performance liquid chromatography
rt.	room temperature
SDS	sodium dodecylsulfate
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
THF	tetrahydrofuran
TLC	thin layer chromatography
TOCSY	total correlation spectroscopy
TMS	tetramethylsilane
p-TSA	*para*-toluene sulfonic acid
[θ]	molar ellipticity (degrees cm^2 $dmol^{-1}$)

Amino acid codes



Amino acid	*Abbreviation*	*1-letter code*	*-R*
Alanine	Ala	A	$-CH_3$
Arginine	Arg	R	$-(CH_2)_3NHC(=NH)NH_2$
Asparagine	Asn	N	$-CH_2CONH_2$
Aspartic acid	Asp	D	$-CH_2CO_2H$
Cysteine	Cys	C	$-CH_2SH$
Glutamine	Gln	Q	$-(CH_2)_2CONH_2$
Glutamic acid	Glu	E	$-(CH_2)_2CO_2H$
Glycine	Gly	G	-H
Histidine	His	H	$-CH_2$(4-imidazolyl)
Isoleucine	Ile	I	$-CH(CH_3)CH_2CH_3$
Leucine	Leu	L	$-CH_2CH(CH_3)_2$
Lysine	Lys	K	$-(CH_2)_4NH_2$
Methionine	Met	M	$-(CH_2)_2SCH_3$
Phenylalanine	Phe	F	$-CH_2Ph$
Serine	Ser	S	$-CH_2OH$
Threonine	Thr	T	$-CH(CH_3)OH$
Tryptophan	Trp	W	$-CH_2$(3-indolyl)
Tyrosine	Tyr	Y	$-CH_2$(4-hydroxyphenyl)
Valine	Val	V	$-CH(CH_3)_2$
Proline	Pro	P	 †

† Complete structure of proline depicted due to cyclic nature

List of Figures

List of Schemes

List of Tables

page

Table of Contents

CHAPTER 1: INTRODUCTION

1.1 Viruses

1.1.1 Introduction

A virus is an infectious agent of small size and simple composition which can multiply only in the living cells of animals, plants, or bacteria. Outside a living cell a virus is a dormant particle, but within an appropriate host cell it becomes an active entity, capable of subverting the cell's metabolic machinery for the production of new virus particles. Historical descriptions of viral diseases date as far back as the 10th century BC. The concept of the virus, however, was not established until the last decade of the 19th century, when several researchers gained evidence that agents far smaller than bacteria were capable of causing infectious diseases.

The question of whether viruses are actually micro-organisms (similar to very tiny bacteria) was resolved in 1935, when the virus responsible for causing mosaic disease in tobacco was isolated and crystallized; the fact that it could be crystallized proved that the virus was not a cellular organism. A virus consists of a single or double stranded nucleic acid core surrounded by a protein capsid. The capsid protects the nucleic acid from enzymic attack and mechanical breakage, delivering it intact to a susceptible host. In some of the more complex viruses, the capsid is surrounded by an envelope that contains fatty material (lipids) and glycoprotein. The complete extracellular product of virus multiplication is called a virion or virus particle.

As well as the importance of understanding the viral infections of animals and man, there is much current interest in viruses for several other reasons. First, viruses are capable of adapting themselves over time and are therefore sources of insight into evolutionary processes as well as the molecular aspects of host-parasite relations. Second, some viruses cause cancers in animals and the possible role of viruses in human cancers is being intensively studied.[1] Third, viral multiplication is a model for cellular development because it requires the sequential expression of genes and the assembly of macromolecules to form highly ordered structures.

Viruses are classified according to the host that they infect: plant, animal or bacteria. Within these three main classes, viruses are grouped according to their chemical composition (type of nucleic acid, presence or absence of a lipid envelope) and their size and shape.

1.1.2 RNA viruses

One of the most striking discoveries made by virologists this century was that the extreme volatility and rapid evolution of some viruses results from the storage of their genetic codes in RNA rather than DNA - a phenomenon not found in any other living organisms.

Because their genetic material is RNA instead of DNA, RNA viruses are of great biological interest. Clearly, their replication must involve radically different mechanisms from those involved in DNA virus replication. For example, the use of RNA as a template instead of the usual DNA means that enzymes are required which are unlikely to be found in the uninfected cells of the host. Such enzymes therefore have to be virus-encoded.[2] Like DNA viruses, the crucial factor in the successful infection of a cell is the production of viral mRNA.

RNA viruses are divided into two groups depending on the nature of their RNA genome:

1. Positive-strand viruses in which the input RNA genome acts as mRNA; these RNA genomes are infectious and, when applied to cells, the purified RNA can initiate a complete infectious cycle of virus replication.
2. Negative-strand viruses in which the input virus RNA genome is transcribed into mRNA. RNA viruses of this sort contain a transcriptase of RNA-dependent RNA polymerase within their particle and since the transcriptase is essential for infectivity, the purified RNA genome is not infectious.

Foot-and-mouth disease virus (FMDV) is an example of a positive-strand RNA virus.[3] In order to replicate, this type of virus particle has to enter a cell, uncoat and deliver its genome intact to the cellular translation machinery where it behaves as mRNA and protein is produced. A complex pathway of proteolytic processing follows to produce the mature virus particles. The genome also has to act as the template for replication by the virus encoded RNA polymerase so that both viral RNA and virus proteins are produced. New virus particles assemble and are released on cell lysis. FMDV is able to complete this cycle of events in about four hours in tissue culture and in the process may produce 100 000 particles from a single cell.

The principal steps in the replication of positive-strand RNA viruses once inside a host cell are illustrated in Scheme 1.1 and described overleaf.

After uncoating, the RNA genome of the infecting particle functions as virus mRNA and attaches to host cell ribosomes for translation into virus-specified proteins; FMDV RNA is polycistronic, *i.e.* it codes for a single very large polyprotein.[3]

Translation: The viral mRNA is translated into a single very large polyprotein which is divided into structural and non-structural proteins. These are produced from the large protein by cleavage due to the action of virus-encoded proteolytic enzymes. The cleavage is a complex process and takes place in several stages.[4] The proteins finally produced include RNA-dependent RNA polymerase and four capsid proteins.

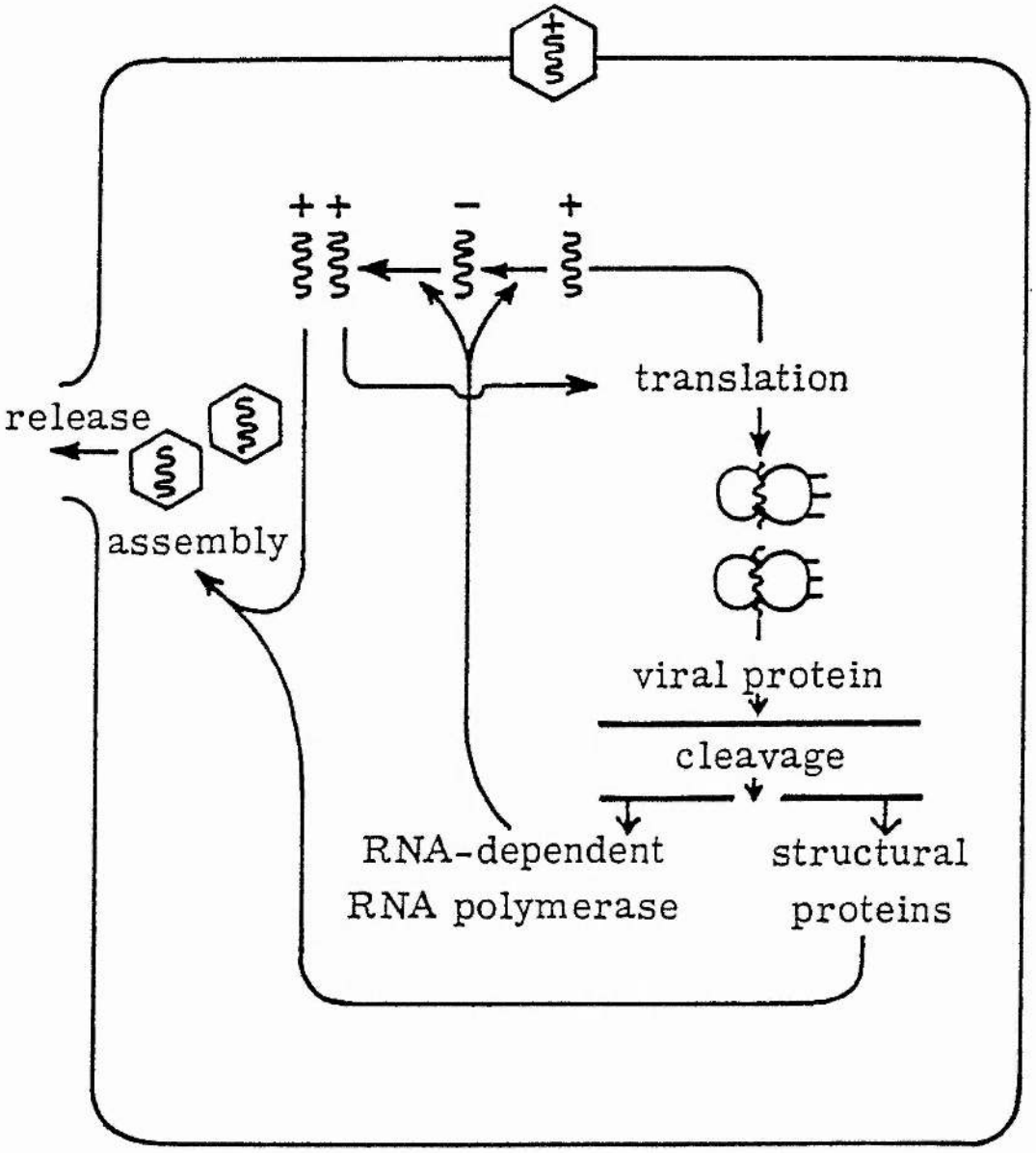
Virus RNA synthesis: New RNA molecules are replicated by a virus-encoded enzyme: RNA-dependent RNA polymerase. This enzyme is produced by cleavage from the large polypeptide which is the primary product of translation. Virus RNA synthesis takes place in the following stages:

1. A second (negative) RNA strand, which is complementary to the input positive RNA genome, is produced with formation of a double-stranded RNA structure known as the replicative form.
2. New positive RNA strands are synthesized, using as a template the complementary negative RNA strand of the double-stranded replicative form.

The new (positive strand) RNA molecules have three functions:

1. Templates for the manufacture of more replicative forms for further synthesis of RNA genomes.
2. Genomes for new virus particles: RNA molecules which become genomes in new virus particles have a small protein linked to their 5′ terminal.
3. Virus mRNA: the newly synthesized RNA which acts as mRNA does not become encapsidated as the genomes of new particles.

Assembly: After replication, the new virus particles are assembled on clusters of ribosomes in the cytoplasm. The new particles are released by sudden rupture of the cell.



Scheme 1.1: *Life cycle of a positive strand RNA virus*

1.1.3 Picornaviruses

One of the largest groups of animal viruses is the picornaviruses which lists among its members a diverse variety of pathogenic agents. The name comes from the Greek *pico* which means tiny and RNA which is the form in which the viruses store their genetic information.

Picornaviruses are non-enveloped and have a 25-30 nm capsid of icosahedral symmetry. They have been subdivided, on the basis of chemical properties such as pH stability and buoyant density, into four main genera: entero-, rhino-, cardio- and aphthoviruses. The known picornaviruses are listed below, grouped into the four main classes.

Table 1.1: *Classification of pircornaviruses*

Genera	*Virus*
❑ Aphthoviruses	FMDV (foot-and-mouth disease virus)
❑ Cardioviruses	EMCV (encephalomyocarditis virus) Mengovirus TMEV (Theiler's murine encephalomyelitis virus)
❑ Enteroviruses	Poliovirus Coxsackievirus Echovirus Enterovirus 70 Swine vesicular disease virus Bovine enterovirus Hepatitis A virus
❑ Rhinoviruses	Human rhinovirus

Each type of virus has a number of different strains or serotypes - for example, there are seven serotypes of FMDV which constitute the aphthoviruses and there are over 100 serotypes of the human rhinovirus (the major cause of the common cold). Altogether, there are over 230 picornavirus serotypes.[3]

In spite of the disparate afflictions caused by these agents, all picornaviruses share remarkable similarity in their genome and capsid organisation.[5] Complete or partial nucleotide sequences are now available for more than 90 strains of picornaviruses, including representative members from all classification groups.[3] Also, the three-dimensional crystal structures have been solved for several of the most interesting types - that of FMDV was published in 1989 (Fig. 1.1).[6]

Despite the modest protein capsid structures and small RNA genomes that characterize this family, picornavirus-caused diseases are of enormous medical and agricultural importance. Infection with rhinovirus for example, better known as the common cold, is among the most prevalent and acute of human upper-respiratory illnesses. Each year, this virus exacts a high toll from every economic and social institution, as cold-related absenteeism is the single largest cause of lost work days for schools and businesses.

Poliovirus, another well known member of the picornavirus family, has likewise afflicted human population for as long as recorded history. Even today, in the presence of modern vaccines and sanitation technology, many thousands of new cases are chronicled annually, especially in developing countries. The unfortunate results produce economic hardships from loss of productive manpower and from the obviously difficult medical burdens incurred in patient care.

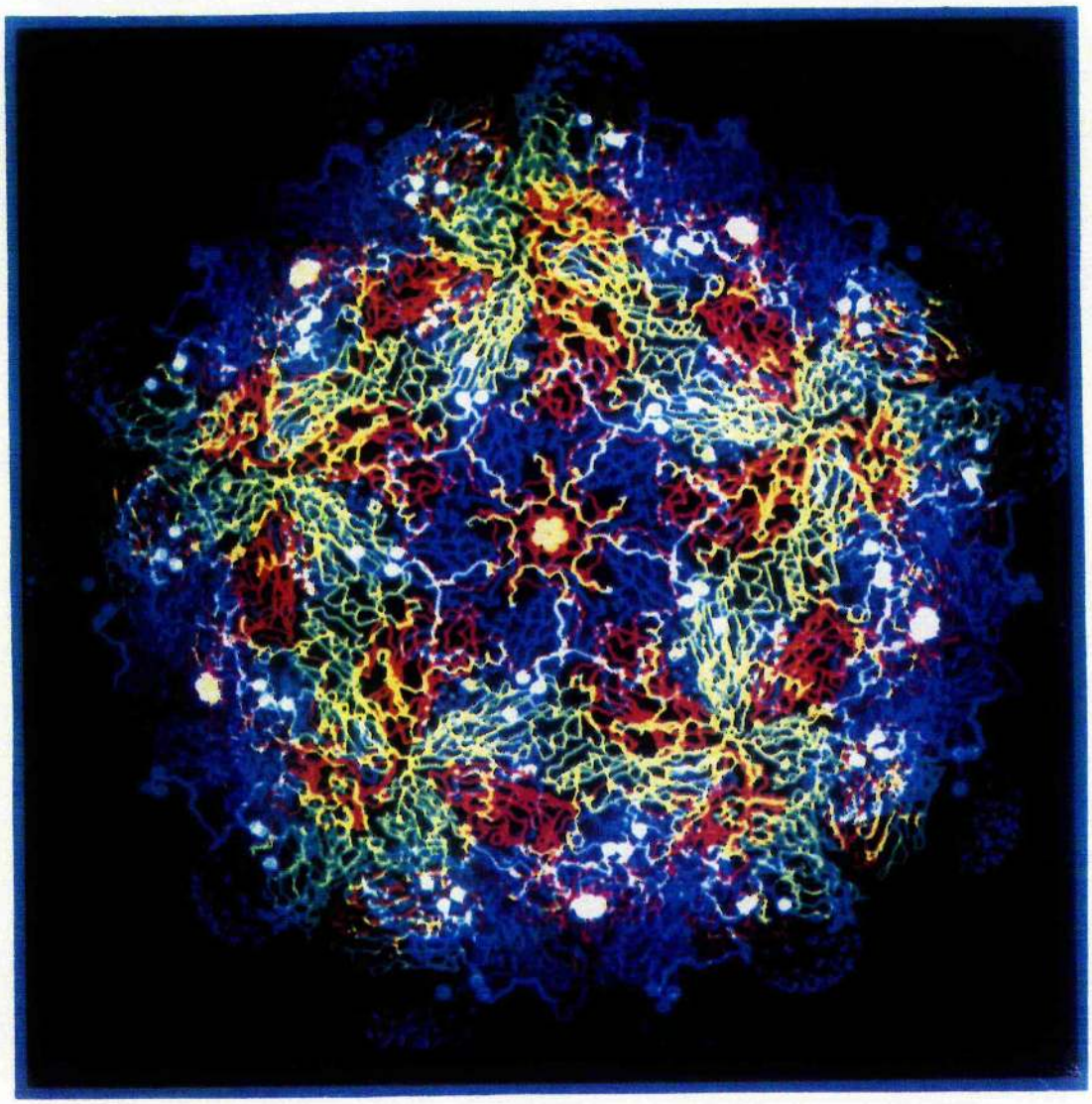


Figure 1.1: *Crystal structure of FMDV*[6]

1.1.4 Foot-and-mouth disease virus

Animal populations are also susceptible to picornavirus infections. Foot-and-mouth disease virus (FMDV) is one of the most virulent pathogens ever described. It affects practically all cloven-footed mammals, including many varieties of commercial livestock, namely pigs, sheep, goats and cattle. The disease is characterized by the formation of painful fluid-filled blisters on the tongue, lips, and on other parts of the body where the skin is thin, such as between the two toes of the feet.

The losses caused by foot-and-mouth disease are tremendous. Malignant forms of the disease have led to losses of up to 50 percent and in those animals that survive, great losses in weight occur because the animals cannot eat. The virus can survive for relatively long periods of time in the air, in food, and even in hides, hair and wool, so measures for controlling the disease are necessarily rigorous.

1.1.5 Control of foot-and-mouth disease

In countries such as Great Britain, where the disease arises only as a result of imported infection, the accepted policy is to stamp it out by slaughtering all affected stock and any other animals which have been exposed to the risk of infection. Slaughtered carcasses, plus bedding and other infected material must either be burnt or buried under six feet of earth. The farm must be thoroughly disinfected and cannot be restocked for a further six weeks. In addition to this, restrictions are placed on the movement of all livestock within a ten mile radius, known as the infected area.

Following the disastrous epidemic of 1967-8, which involved 2397 outbreaks and payments in compensation to owners of slaughtered animals of about £27 million, a committee was appointed under the chairmanship of the Duke of Northumberland to review the policy and arrangements for dealing with the disease. The Northumberland committee recommended continuation of the slaughter policy. It further advised reinforcement by a ring vaccination scheme if meat import restrictions were not implemented.

There are several reasons why it is hoped that these measures will never be used. First, once Britain becomes an infected country, it will lose many of its export markets. Second, because there are a variety of different strains of FMDV, the vaccine in use may not be totally effective. Third, vaccinated animals can become

carriers, shedding infection to other stock. Finally, to give full protection, vaccination would have to be carried out each year and in the long term, this would be much more expensive than the current slaughter policy.

Considerable attention is being paid to methods of preventative inoculation and vaccines are widely used on the European mainland, in South America and Africa. Up to the present day, however, no method has been used in Europe which will give an animal immunity to more than one or two of the several strains of virus which can cause the disease. Another disadvantage is that generally, immunity lasts only six months. In countries where the disease is endemic and its incidence high, such as Argentina, Germany, Italy and Spain, comprehensive vaccination is seldom practical for reasons of cost. Instead, ring vaccination of all susceptible animals within a given radius is usually practised. Other countries such as Denmark, Sweden, Switzerland, Holland and Mexico adopt a combined slaughter and vaccination policy.

1.1.6 Vaccines

There is no general useful therapy for viral infections in any way similar to antibiotic therapy for bacterial disease. The vaccines that are currently employed are produced from chemically inactivated forms of the virus. There are several problems with the use of such vaccines.

1. There are a variety of different strains of FMDV and several of these are difficult to grow in sufficient abundance to provide enough material for effective vaccination.
2. The virus particle is unstable, especially under acidic conditions and at warm temperatures. Thus it is necessary to refrigerate vaccines to retain their potency which is clearly a major limitation, particularly in tropical climates.
3. There have been several cases of outbreaks of disease due to improperly inactivated virus.[7]

The concept of using a peptide vaccine stems from work on the tobacco mosaic virus which showed that a segment of the coat protein would elicit an antibody which reacts with the virus particle.[8] Clearly, the use of synthetic peptides as therapeutic agents would circumvent the hazards of working with large scale production of the infectious virus.

Of the four capsid proteins of FMDV, (VP1-VP4), only VP1 has been shown to have an important antigenic role.[9] VP1 was found to be the only capsid protein which induced neutralizing antibody responses against FMDV.[10] Treatment of the virus particle with trypsin, which causes the cleavage of VP1 only, results in a drastic reduction of the infectivity of the virus.[11,12] Several trypsin-sensitive regions within VP1 have antigenic activity, the most significant being the region bounded by residues 130 and 160 with a minor contribution from residues 200 to 213 at the extreme carboxyl terminus. The major immunogenic site on VP1 (130-160) has been structurally characterized in detail using X-ray crystallography[13] and consists of a mobile surface loop connecting two β-strands (Fig. 1.2). The residues Arg-Gly-Asp (145-147) are known to be particularly important as a recognition site for cell attachment and in VP1, the RGD sequence is highly conserved and centrally placed in the loop as shown in the diagram below. The same triplet has been located in an external loop of the coat protein VP1 of the human rhinovirus[5] and it has been proposed that this unit is involved in the binding of the virus particle to the cellular receptor.[14,15] Further work is required to establish the identity of the cellular receptor for FMDV and to ascertain the nature of the interaction of the virus particle with it.

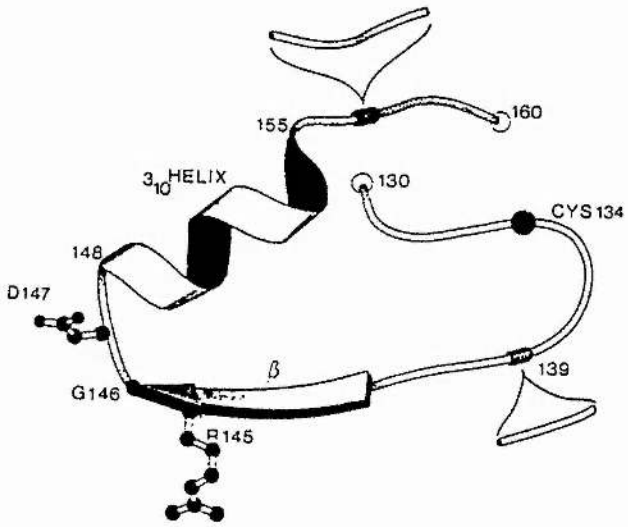


Figure 1.2: *Major immunogenic loop of FMDV*

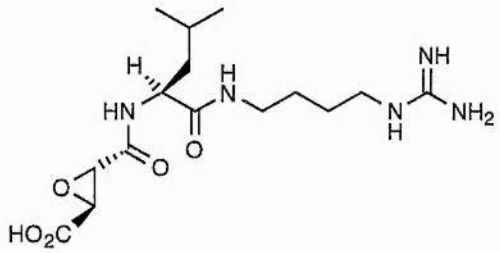
Chemically synthesized peptides corresponding to two different regions of the VP1 protein (Y130-P160 and R200-L213) produced significant levels of serotype-specific virus neutralizing antibodies in cattle, guinea pigs and rabbits.[16] Protection of cattle against FMDV was achieved after a single immunization with a 40 residue peptide which consisted essentially of the two previously identified immunogenic regions connected by a diproline spacer.[10] Various experiments were carried out with a range of peptides containing either or both of the two significant sequences.[17] These experiments showed that the most effective structure, in terms of protection is one in which both sequences are present with no additional amino acids. Although the sequence Y130 to P160 appears to be the major site for generation of neutralizing antibodies, the additional presence of the sequence R200-L213 significantly increases the potency of the peptides.

One of the disappointing aspects of FMDV peptide vaccines has been their relatively poor performance in cattle compared with the high promise shown in guinea pigs. It has been suggested that to realise their potential as vaccines, the peptides must contain domains which react with helper T-cell receptors in addition to binding sites for anti-protein antibodies. Preliminary experiments have shown that genetically controlled non-responsiveness to a peptide can be overcome by adding an additional helper T-cell epitope from a different protein. Mice, which are non-responders to immunogenic peptides of FMDV will respond to the same sequences when they are linked to helper T-cell determinants from sperm whale myoglobin.[18]

Two very recent studies have used novel approaches, which do not focus on VP1 as a therapeutic target, to try to disable the virus. Lewis *et al.* have demonstrated that empty virus capsids isolated from infected cells antigenically resemble virus particles.[19] They have constructed FMDV clones which can express, in an appropriate host, conformationally correct immunogens which lack nucleic acid and could make effective and safe vaccines. Plasmids containing the FMDV structural protein precursor P1 and the 3C proteinase genes were expressed in *E. coli*. This resulted in the efficient synthesis and processing of P1 and assembly of the resultant proteins into empty capsids. This material elicited a significant neutralizing antibody response in vaccinated guinea pigs but has yet to be tested in cattle.

The FMDV leader proteinase L, possesses considerable potential as a target of antiviral drug therapy.[20] Kleina and Grubman have shown that the thiol proteinase inhibitor E-64 (S-*trans*-epoxysuccinyl-leucylamido-(4-guanidino)-butane) (**1**) specifically blocks autocatalytic activity of the proteinase L and interferes with the cleavage of the structural protein precursor P1.[21] L proteinase activity is specifically

and irreversibly inhibited by E-64 and its analogues, and inhibition of the L proteinase initiates a cascade of events that affect the replication of FMDV. Processing of the structural protein precursor P1 is blocked, leading to improper assembly of the viral capsid.[22] The L proteinase is an ideal candidate molecule for antiviral compounds as among the serotypes of FMDV, L proteinases share greater than 90% homology at their amino acid level. E-64 blocks the autocatalytic activity of the L proteinase in all serotypes tested. Thus, an antiviral agent that attacks L would circumvent the problems of serotype-specific reactions observed with conventional FMDV vaccines.



(**1**): *Thiol proteinase inhibitor E-64*

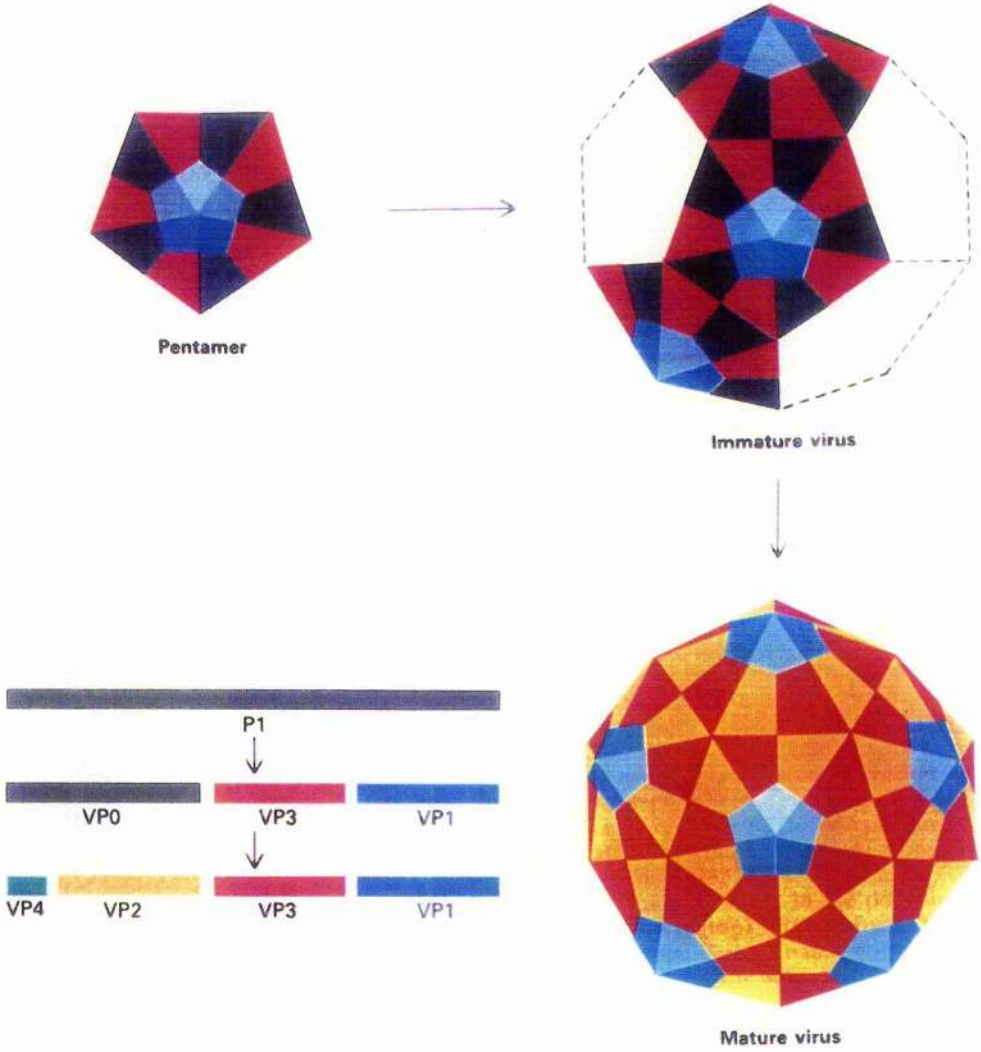
1.1.7 Structure of FMDV

The virus particle (25-30 nm in diameter) consists of a single copy of the genome packaged in an icosahedrally symmetric protein capsid composed of 60 copies of the four structural proteins, VP1-4.[6] The primary functions of the viral capsid are:

1. To interact with specific cellular receptors on susceptible host cells in a manner that results in the internalization of the genome.
2. To selectively package the newly synthesized genome.
3. To facilitate the eventual exit of the new virus particle.

The icosahedral capsid is assembled from pentameric intermediates, each composed of five copies of VP1, VP3 and a precursor protein, VP0, in which VP4 and VP2 are covalently linked (Scheme 1.2).[23] Cleavage of VP0, one of the final steps of capsid assembly, may be related to encapsulation of the nucleic acid and stabilization of the mature particle.[24,25] At the surface of the virus, the area surrounding the fivefold axis of the pentamer is composed of the main capsid proteins, VP1-3. VP4 is much smaller, has an extended conformation and is located internally underneath the other proteins.

One or two copies of VP0 remain uncleaved in each capsid, the functional significance of this observation being unclear.[26] The pentamers appear to be stabilized by interactions between the N and C termini of VP1 and VP3 together with VP4. Adjacent pentamers are held together by hydrogen bonds between parts of VP2 and VP3 and it is thought that the weakness of these interactions may be important in the uncoating of the viruses.[3] The icosahedral shell surrounds the RNA genome which is linked covalently to a small virus-encoded protein, VP_g.[27] Upon entry into the cell and following uncoating, this RNA is translated into one long polyprotein which is cleaved by at least two viral proteinases to give the structural and non-structural proteins.[4,28]



Scheme 1.2: *Assembly and maturation of the FMDV virion*

The major portion of the virus genome is a single very large open reading frame of 6996 nucleotides encoding a polyprotein of 2332 amino acids.[29,30] The polyprotein may be considered as four different components, termed L, P1-2A, P2 and P3. The P1-2A section is the precursor of the capsid proteins while the P2 and P3 precursors are processed to non-structural proteins involved in virus RNA replication and protein processing.

The structure of the foot-and-mouth disease virus has been determined at close to atomic resolution by X-ray diffraction (Fig. 1.1).[6] Unlike other picornaviruses, FMDV does not possess any major depressions or pits on its surface and as VP1-3 are substantially smaller, the protein shell is generally thinner.[5] The smoothness of the surface has important implications for cell attachment and it is thought that FMDV binds to cellular receptors *via* a small, flexible exterior loop on VP1 (Fig. 1.2). Another unique property of FMDV, which is not demonstrated by any of the other picornaviruses is extreme susceptibility to acid induced dissociation. It has been suggested that the high number of histidine residues lining the pentamer interfaces is responsible for the instability of the virus below pH 7.[15] It may be that the acid sensitivity of the virus is a necessary prerequisite for rapid and efficient uncoating within the cell.

The structural proteins VP1-3 consist of eight strands of β-sheet arranged as a sandwich of two four-stranded sheets. The tertiary structure is formed by a β-hairpin bend which results in a wedge shaped molecule which fits well around the five-fold axes of the icosahedral capsid (Fig. 1.3).

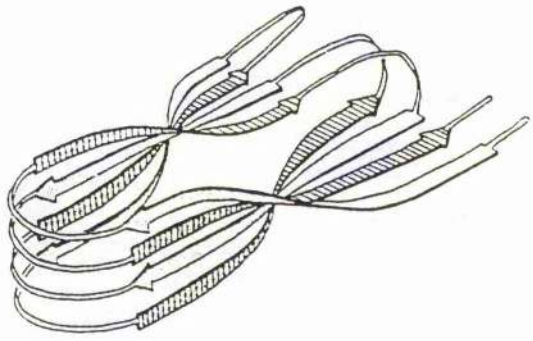


Figure 1.3: *Structure exhibited by the three external capsid proteins*

1.1.8 Polyprotein processing

In the life-cycle of all picornaviruses, the RNA genome is translated into a very long polyprotein which undergoes a complicated series of proteolytic processing events to produce the mature virus particles. Several excellent reviews have been published on picornaviral polyprotein processing.[3,4,31,32] The complete polyprotein is never observed within infected cells since some of the processing events occur extremely rapidly, probably while the translation process is still occurring.

A three-tiered cascade of primary, secondary and maturation cleavages efficiently produces the spectrum of viral proteins necessary for a productive infection. All the active proteins associated with picornaviral processing (2A, L, 3C and VP0) are encoded by viral genomes and no cellular components are required. Enzymic self-sufficiency must clearly be an advantage to a highly mobile pathogen and may explain in part why the picornaviruses can efficiently infect a diverse variety of hosts with such virulence and obvious success.
The three phases of viral processing are described below.

1. **Primary cleavage.**
   These processing events are cotranslational, occurring as the ribosome moves along the genome. In FMDV, the polyprotein is rapidly processed into four primary cleavage products: L, P1-2A, 2BC and P3 as shown in Fig. 1.4. These primary cleavages involve all three viral proteinases: L, 2A and 3C.

2. **Secondary cleavage.**
   This phase involves the subsequent processing of the primary cleavage products into all the mature viral proteins necessary for establishment and completion of a successful infectious cycle. An example is the processing of the precursor protein P1 by the viral proteinase 3C to give the capsid proteins (VP1-VP4 or 1A-1D). The details of 2BC and P3 secondary processing have not been fully characterized in FMDV or other picornaviruses but it is known that 3C is solely responsible for all secondary processing reactions.[20,33,34]

3. **Maturation cleavage.**
   This is the final cleavage within picornaviral polyproteins and involves the processing of the VP0 structural protein into VP2 and VP4. The cleavage only occurs on encapsulation of viral RNA to produce virus particles (Scheme 1.2) and the mechanism is so far unknown.[23,24]

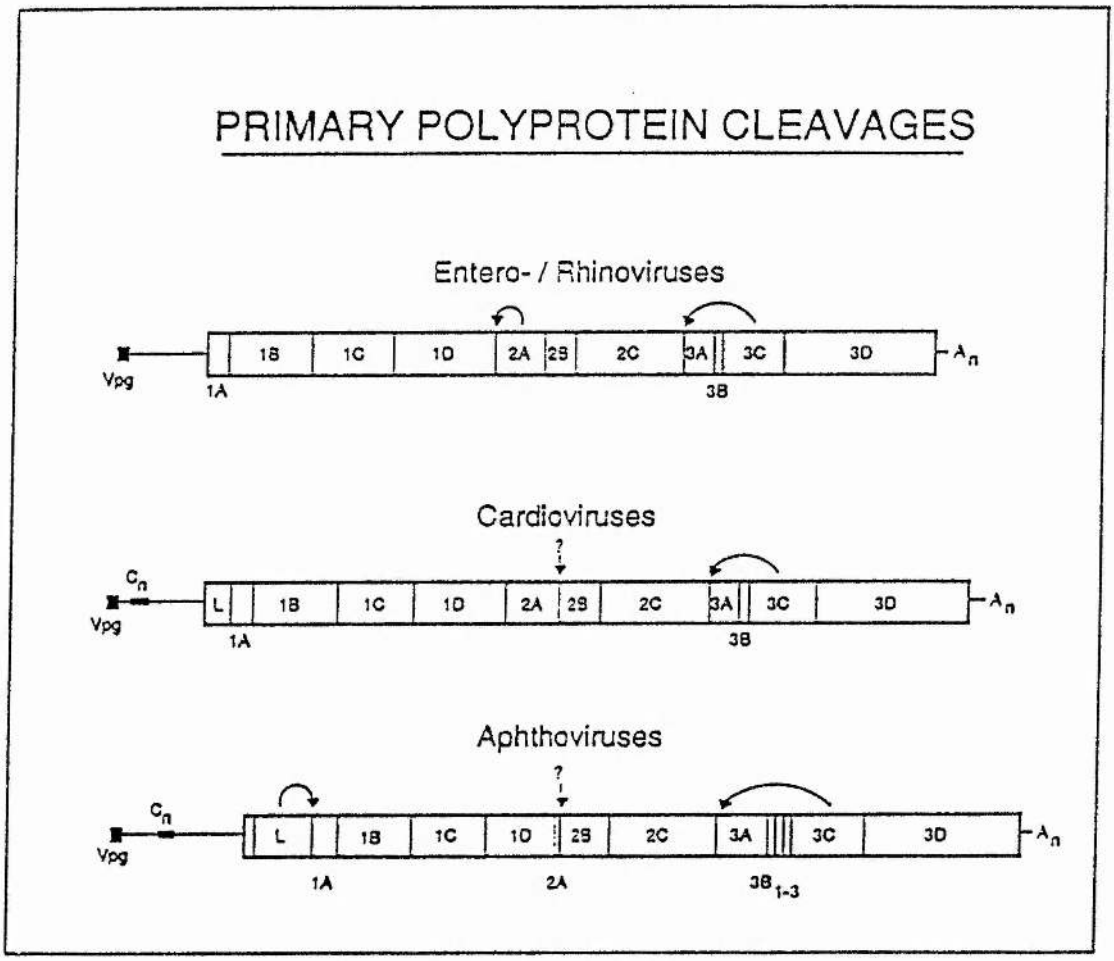


Figure 1.4: *Picornavirus primary polyprotein cleavages*

1.1.9 Primary polyprotein cleavages

There are three primary processing events which occur within in the FMDV polyprotein to produce the primary cleavage products: L, P1, P2 and P3.

1. **The L/P1 cleavage**
   The FMDV L protein is located at the N-terminus of the polyprotein and is the first gene product released from the nascent polyprotein.[35] Thus, in FMDV, the first processing event is uniquely the L/P1 cleavage. In cardioviruses, the only other picornaviruses which have an L protein, the L/P1 junction is cleaved by the 3C proteinase.[14]

2. **The 2A/2B cleavage**
   The second very rapid cleavage event within the FMDV polyprotein occurs at the 2A/2B junction. In contrast to entero- and rhinoviruses, the cleavage in cardio- and aphthovirus polyproteins occurs at the carboxy terminus of their 2A regions, releasing a P1-2A precursor by autocatalytic cleavage. This junction always appears completely cleaved, even using *in vitro* translation assays[33,36] and the cleavage has been shown to be independent of both L and 3C.[36] In cardioviruses, the 2A/2B cleavage is attributed to the 2A proteinase which contains 143 amino acids.[32] In FMDV, a similar cleavage occurs but the 2A region is only 16 amino acids in length.[30] Studies on cardioviruses have shown the 2A/2B cleavage to possess characteristics of a cotranslational intramolecular cleavage: insensitivity to dilution and resistance to a range of proteinase inhibitors.[37]

3. **The 2C/3A cleavage**
   The P3 region of the polyprotein contains the 3C proteinase which mediates not only the 2C/3A primary cleavage but all subsequent secondary cleavages. The mechanism of the 2C/3A cleavage is not fully understood but it is known to be an intramolecular cleavage for which the 3C proteinase is solely responsible.

1.1.10 Viral proteinases

Proteinases play a key role in the replication cycles of many viruses.[2] One or more proteinases are usually involved and can be of host or viral origin. Host proteinases are frequently involved in the maturation of viral proteins that are eventually

associated with the viral envelope whereas viral proteinases usually function in the cytoplasm of the host cell early in the replication cycle.[4]

Virus-encoded proteinases appear to fulfil one of two functions during the replication cycle. First, most are involved in the processing of a high-molecular weight polyprotein into functional gene products. The second general process involves the maturation of structural proteins during the virion assembly process.

In viral proteolytic processing, two types of cleavage events are observed: a *cis* or monomolecular reaction and a *trans* or bimolecular reaction. In cases of *cis* cleavage, the reaction is intramolecular, dilution independent and follows zero-order kinetics. The *cis* reactions, which are often quite rapid and occur during translation are also referred to as 'autocatalytic'. The *trans* reaction is exemplified by a protein with proteolytic activity cleaving a second molecule which contains the substrate cleavage site. This is an intermolecular reaction which proceeds with second-order kinetics and is sensitive to dilution.

In FMDV, there are three identified virus-encoded proteinases. These are described below and discussed in relation to their counterparts in other picornaviruses.

1. **The L proteinase**

   Amongst the picornaviruses, only the cardioviruses and FMDV have an L or 'leader' protein, situated at the polyprotein N terminus and only in FMDV is the L protein a proteinase.[38] The precise nature of L is not known but recently, a thiol proteinase inhibitor has been shown to inhibit its function in FMDV, suggesting that L is a cysteine-type proteinase.[21] The main function of L in aphthoviruses is thought to be the suppression of host protein synthesis in infected cells.[20] This is achieved by inactivation of the cap-binding protein complex which is required for the formation of ribosome initiation complexes with capped mRNAs. The L proteinase proteolytically cleaves p220, a component of the cap-binding protein complex resulting in the shutoff of host cell translation during infection by FMDV.[22]

2. **The 3C proteinase**

   All picornaviruses use the activity of viral polypeptide 3C as their major source of proteolytic cleavage activity.[4] As all 3C processing reactions are sensitive to thiol inhibitors, the proteinase was originally categorized as a cysteine-type agent.[31] Subsequently, sequence comparisons have shown that picornavirus 3Cs are

probably more closely related to serine proteinases and bear a strong resemblance to trypsin and chymotrypsin.[2,26] Thus, it seems that the 3C proteinases, although containing a nucleophilic thiol residue, are structurally similar to certain cellular serine proteinases. X-ray crystallographic studies are needed to show the true structure of the active site. All picornavirus 3C proteinases have a high substrate specificity and almost exclusively catalyse the proteolysis of viral precursor polyproteins.[20,39] In addition to strict substrate requirements for proteolytic cleavage, the 3C proteinases have interesting properties as enzymes, participating in both uni- and bimolecular reactions.[40] EMCV 3C has been shown to cleave itself from precursors in an autocatalytic manner and also to participate in the cleavage of substrates in dilution-dependent bimolecular reactions.[37]

3. The 2A proteinase

Cleavage by the 2A proteinase in aphtho- and cardioviruses is fundamentally different from that of the rhino- and enteroviruses. The FMDV 2A proteinase cleaves at its C terminus to generate a P1-2A intermediate. In contrast, the 2A proteinase of the rhino and enteroviruses performs a *cis* cleavage at its own N terminus to separate the P1 precursor from the growing 2A-P2-P3 polyprotein. Interestingly, the 2A proteinase of rhino- and enteroviruses also has an important *trans* cleavage activity, indirectly inducing cleavage of the p220 component of the cap-binding protein complex through activation of latent cellular proteinases.[41] This results in selective inhibition of host cell protein synthesis. Thus, the rhino- and enterovirus 2A proteinases are performing the function for which aphthoviruses have evolved a separate proteinase, L (see above). It should be noted that for cardioviruses, neither the L protein nor the 2A proteinase have shown any activity which would lead to cleavage of p220.

On the basis of sequence alignment and site-directed mutagenesis studies, rhino- and enterovirus 2A proteinases have been shown to resemble trypsin-like serine proteinases but with an active-site cysteine nucleophile instead of serine.[42,43] The activity of the cardio- and aphthovirus 2A proteinases has not yet been resolved. Although cardiovirus 2A proteins contain 143 residues, the FMDV 2A region is only 16 amino acids in length and it is unprecedented for such a short peptide sequence to have the ability to mediate autocatalytic cleavage. It has not been conclusively proven whether the 2A/2B junction is very efficiently recognized by host proteinases in a wide variety of host cells or whether the 16 amino acids represent a very small self-cleaving proteinase.

1.1.11 FMDV 2A: active element or passive substrate?

There are three explanations which may account for a cotranslational cleavage event occurring with such a short amino acid sequence.

1. FMDV 2A functions as a substrate for a cellular proteinase which, to account for the observed cleavage kinetics, would need to be closely coupled to translation.

2. The FMDV 2A sequence in some manner disturbs the normal peptide bond formation during translation.

3. The FMDV 2A sequence possesses an entirely novel type of proteolytic activity.

Various investigations have provided evidence which is consistent with the third possibility.

- FMDV 2A shows high sequence similarity to the C terminal region of the much longer 2A protein of cardioviruses. Molecular biology studies have shown that when the amino-terminal two-thirds of EMCV 2A (Table 1.1) are deleted, this does not block its proteolytic activity.[43]

- Cleavage at the FMDV 2A/2B junction shows characteristics of a unimolecular event *e.g.* insensitivity to dilution and resistance to a range of proteinase inhibitors.[32,36]

- The cleavage occurs in a range of heterologous expression systems: rabbit reticulocytes, human HTK-143 cells, wheatgerm extracts and insect cells.[36,44]

- Recent studies have shown that introduction of a 16-19 amino acid segment, including the FMDV 2A sequence, into a totally foreign protein induces cleavage of the novel protein at the 2A/2B junction.[44,45]

Although the mechanism of FMDV 2A mediated cleavage is not yet understood, its utility is apparent. The strategy of self-processing polyproteins may be adapted in many applications to enable the coordinated and stoichiometric expression of multiple proteins from a single open reading frame.

1.2 Peptide structure

1.2.1 Protein and peptide conformation

One striking characteristic of proteins is that they have well-defined three-dimensional structures. A denatured protein, which consists of a randomly arranged polypeptide chain is devoid of biological activity. The biological function of a protein is, therefore, dependent on its conformation, which is the three-dimensional arrangement of its atoms.

Small peptide fragments of proteins have also been shown to adopt relatively stable, well-defined structures in both organic solvents and aqueous solution.[46-63] In general, short linear peptides are more likely to exhibit structure in organic solvents than in aqueous solution.[61,64] Common structural features adopted by linear peptides are β-turns and helices which are most often detected in solvents such as dimethylsulfoxide (DMSO), trifluoroethanol (TFE) and methanol. The ability of membranes to influence the conformation of short peptides has also been illustrated in several instances.[62]

Since the FMDV 2A polypeptide has been shown to be active in foreign proteins, it can be assumed that its active conformation is not structurally dependent on other parts of the FMDV polyprotein.[65,66] The 2A region is probably not large enough to contain tertiary structural elements so one must consider its three-dimensional conformation in terms of secondary structural units such as turns, helices and β-sheets.

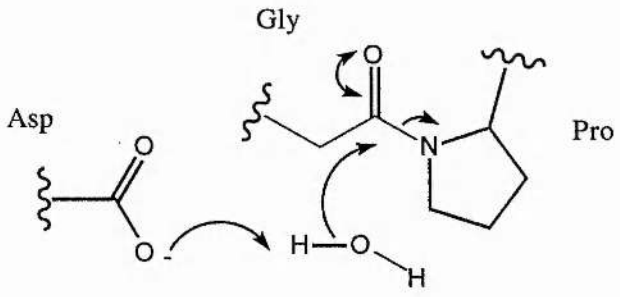
1.2.2 Conformation and mechanism of 2A

The self-cleavage hypothesis involves cleavage as an intrinsic property of the 2A region, which may form a structure that is unstable under physiological conditions. Computer modelling studies and experimental results (see chapter 2) have indicated that a favourable conformation for the 2A peptide would contain an α-helical region and also a reverse turn in the carboxyl terminal section. Calculating the most energetically favourable conformation of the 2A peptide will provide strong clues as to its mechanism of action.

If the 2A peptide is indeed a small, self-cleaving proteinase, then it must have a suitable nucleophile positioned close to the Gly-Pro scissile bond.

Many groups in active sites of enzymes serve as acid or base catalysts acting upon different groups in a substrate and thereby contributing to rate acceleration. Concerted general acid or base catalysis is particularly effective. Several amino acid side chains can act as general acid or base catalysts in enzymes, including those of glutamic acid, aspartic acid, lysine, tyrosine, histidine, serine and cysteine. In their protonated forms they are acid catalysts, and in their unprotonated forms they are base catalysts.

The FMDV 2A region does not contain any cysteine, tyrosine or histidine residues. It contains one serine residue but from mutation experiments, it is not thought to be involved directly in catalysis.[67] Site-directed mutagenesis studies have also shown that the Glu, Asp, Lys and Asn residues in 2A are essential for cleavage.[67] Since only one of these is likely to act as a nucleophile in the cleavage mechanism, the others must all be important from a structural point of view. Molecular modelling studies[67] have shown that in one energetically favourable conformation, the aspartate residue could be situated close to the scissile bond, with a strong hydrogen bond existing between the Glu and Asn residues (Fig. 1.5). Therefore, a possible mechanism for the mode of action of FMDV 2A is one where an acidic residue, such as Asp, participates in a general base catalyzed hydrolysis of the Gly-Pro bond (Scheme 1.3).



Scheme 1.3: *Possible mechanism of cleavage of the FMDV 2A peptide*

The aspartate residue can only be situated sufficiently close to the cleavage site if the Gly-Pro amide bond is in the *cis* conformation. This causes a sharp β-bend in the peptide backbone. An optimised structure of the 2A region in a potentially cleavable conformation (calculated in a low dielectric medium) with a *cis* Gly-Pro bond is shown overleaf (Fig. 1.5).

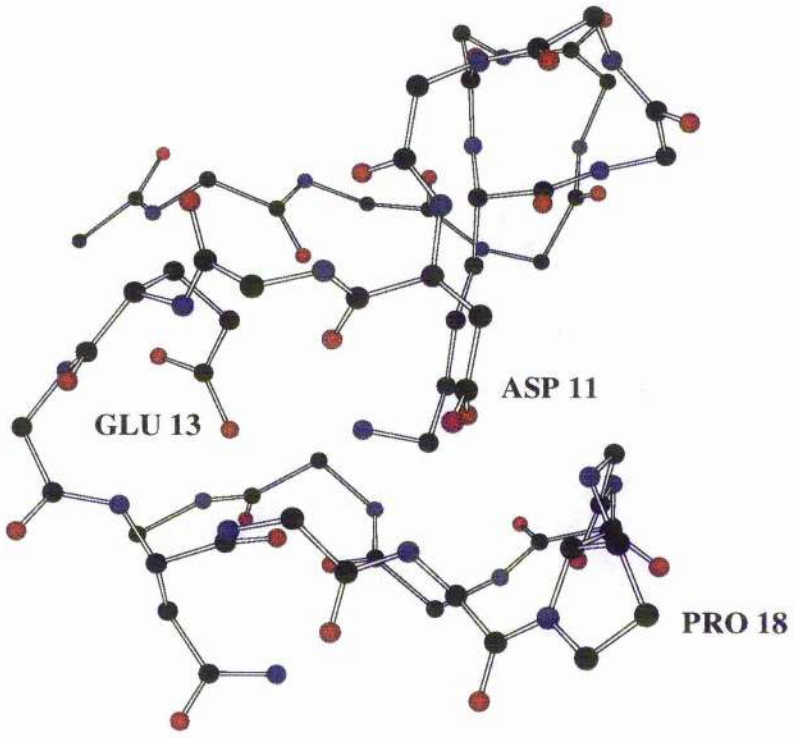


Figure 1.5: *Optimised structure of the 2A region in a low dielectric medium*

Molecular biology experiments have been carried out in which a series of recombinant polyproteins were constructed by deleting either the upstream or downstream context of the 2A region.[67] Analysis of the processing properties of these recombinant polyproteins showed than the sequence of the region downstream of the N-terminal Pro of 2B did not affect cleavage activity and the upstream context, although influential in, was not crucial for activity. Other deletion experiments showed that truncated versions of the 2A region in which the first few N-terminal residues were removed could also mediate cleavage.[67]

In the FMDV polyprotein 2A region, the sequence at the cleavage site is Pro-Gly-Pro with scission occurring between the glycine and the second proline. The presence of a proline residue in a peptide or protein is expected to result in unique conformational preferences that may, or may not, be biologically significant.[68] In globular proteins, most proline residues occur at protein-aqueous interfaces, where they are involved in β-turns. Only 12% of prolines were found to be buried in protein interiors, compared with 38% for an average amino acid.[69]

1.2.3 Structural influence of proline

The influence of proline on secondary structure can be summarised into three main classes.

- Proline is a classical breaker of α-helical structure in proteins. It is hardly ever found at the carboxyl terminus of a helix and is rarely in the centre of this structure.
- A high proportion of proline residues are found in reverse turns. Over 48% of prolines which are followed by glycine are involved in a turn.[70]
- With proline, the possibility of *cis/trans* isomerism of the X-Pro peptide bond arises. The observed abundancy of *cis* X-Pro bonds in small peptides is between 10 and 30%, as compared to less than 0.1% for other amino acids.[70]

While individual proline residues are tolerated in β-strands and the first turn of helices, proline pairs, whether adjacent of separated by several residues, strongly inhibit the formation of classic secondary structural units. Thus, proline-rich regions of proteins are unlikely to adopt conventional structures.

Proline residues occur in many proteins as random single units, in pairs or included in multiple tandem repeats.[71] Perhaps the most striking of these is the one observed in the circumsporozoite protein of *Plasmodium falciparum* (one of the malarial parasites) where Asn-Ala-Asn-Pro is repeated 37 times.[72,73] Other proline-rich proteins have been isolated and characterized from diverse sources, including bovine colostrum, rat prostate, serum chylomicrons and the respiratory tract, in addition to the saliva of rats, humans and rabbits.

There are very few crystal structures of proline-rich regions in proteins so most of the structural information on such sections has come from NMR and CD spectroscopy as well as predictions from molecular modelling studies.[71] It is thought that the role of proline in such regions is to stabilize extended structures, such as the polyproline II helix, an extended structure with three residues per turn. This is the preferred conformation in solution for a sequence of four or more consecutive prolines and is also commonly adopted by repeat sequences in which every third residue is a proline.

Many viral proteins are now known to contain segments rich in proline repeats, for example, polyoma VP1 protein, simian virus 40 VP1, influenza virus haemaglutinin and hepatitis B core antigen. A nuclear protein in Epstein-Barr virus has no fewer than 29 proline residues in succession.[71,74]

1.2.4 Occurrence of *cis* proline residues

In two recent protein database searches,[70,75] the frequency of *cis* proline residues in globular proteins whose structures have been determined by X-ray crystallography was calculated. The percentage of proline residues that form *cis* peptide bonds was around 6%, as compared with an average of 0.05% for other amino acids. It is a matter of debate as to whether the amino acid sequence in the immediate environment of a proline residue influences the isomeric state of the X-Pro bond.

Residue 111 of the capsid protein VP1 of FMDV is a *cis* proline (see section 1.3) and this residue in not conserved as a proline in any of the other picornaviruses. Residue 84 of VP2 on the other hand, is also a *cis* proline and is conserved across the entire family of picornaviruses.[13]

Using information from a protein crystallography database, the frequency of preceding residues in protein *cis* X-Pro bonds has been recorded (Fig. 1.6).[70] Since only 59 proteins were included in the search, the number of occurrences for every residue is less than ten which is a small sample size. However, the high occurrence of tyrosine may be noted, and glycine also appears to be a rather favourable preceding residue for formation of a *cis* X-Pro bond.

In the FMDV polyprotein 2A region, the sequence at the cleavage site is Pro-Gly-Pro with scission occurring between the glycine and the second proline. Of the 87 Pro-X-Pro entries on the database, the statistical spread of possible conformations is shown in the table below.

Table 1.2: *Conformations adopted by X-P-X motifs in proteins*

trans P - X - *trans* P	80
trans P - X - *cis* P	3
cis P - X - *trans* P	3
cis P - X - *cis* P	1

Thus, the proposed assignment of the terminal section of the FMDV-2A peptide as *trans* P - G - *cis* P as illustrated in Fig. 1.5 would be unusual but certainly not unprecedented. *cis* Proline residues are found primarily in bends and turns (73%), suggesting a specific structural role for this type of bonding.[75]

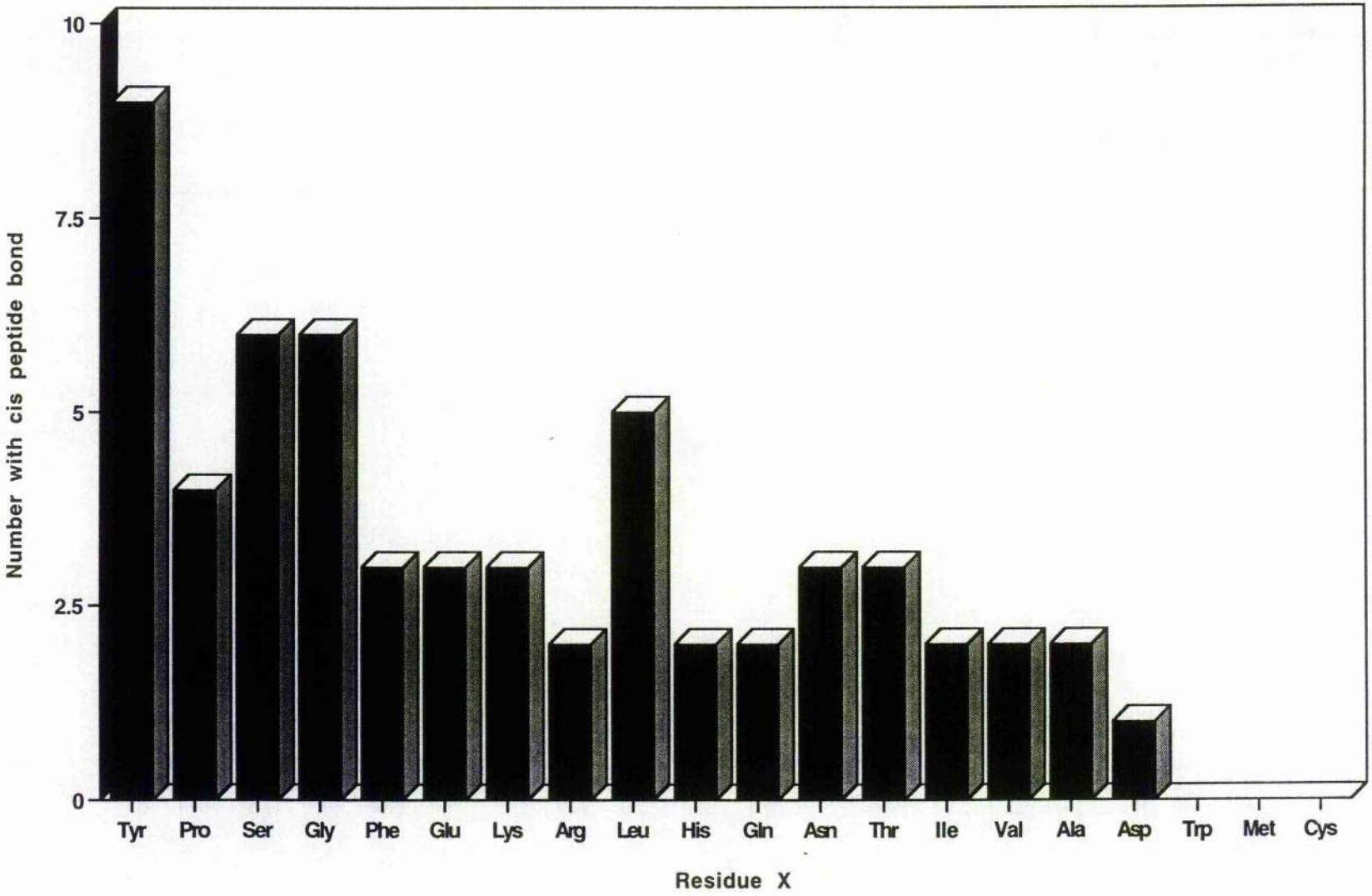


Figure 1.6: *Frequency of* cis *X-Pro bonds from protein database*

1.2.5 Reverse turns

Globular proteins have rather spherical structures, in spite of being composed of straight segments of polypeptide chain. This is because the polypeptide chain generally makes rather sharp bends at the surface, thereby reversing the direction of the chain. It is these reverse turns that give proteins their globular shape, and a substantial fraction of residues of every protein are involved in them. They are also known as hairpin bends, β-bends, and β-turns, because they often connect antiparallel β-strands.

A β-turn is a segment composed of four amino acids (*i* to *i*+3) that occurs when a peptide strand changes direction. If, on the other hand, the segment is composed of three amino acids, one speaks of a γ-turn. The standard criterion for a β-turn is that the distance between the α carbons of the first and fourth residues is less than 7Å.

Although four consecutive residues in the polypeptide chain are generally considered to comprise a reverse turn, it is only the torsion angles of the second and third residues (*i*+1 and *i*+2) that are critical. The first and fourth residues are usually included because a hydrogen bond exists between them in most cases.[76] All observed chain reversals involving four residues can be subdivided into ten types: I, I′, II, II′, III, III′, IV, V, VI and VII.[76,77] These types are characterized by the values of the observed dihedral angles, $\phi(i+1)$, $\psi(i+1)$, $\phi(i+2)$ and $\psi(i+2)$. An example of an 'ideal' type I bend where the four angles are -60°, -30°, -90° and 0° respectively is shown below (Fig. 1.7).

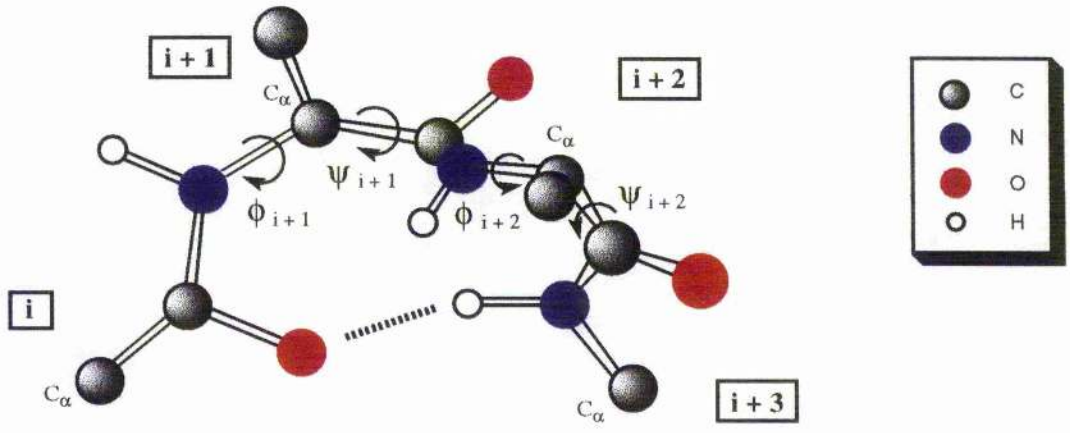


Figure 1.7: *An ideal type I β-turn*

Many examples of turns are observed in proteins, but the conformational angles are not usually ideal, and a hydrogen bond is not always formed between residues *i* and *i*+3. A type VI bend owes its existence to a *cis* peptide group between residue *i*+1 and a proline at position *i*+2 in the bend.[78,79] A new type of structural motif, the pseudo βI turn which also contains a *cis* X-Pro peptide bond has recently been identified in cyclic hexapeptides.[80]

A type VI turn has been observed in the X-ray crystal structure of the protein Bovine Ribonuclease S (Fig. 1.8).[76] The turn occurs at residues 91-94 which are Lys-Tyr-Pro-Asn respectively. The *cis* peptide group is geometrically suited for reversing the direction of the polypeptide chain and no hydrogen bond is needed to stabilize the turn.

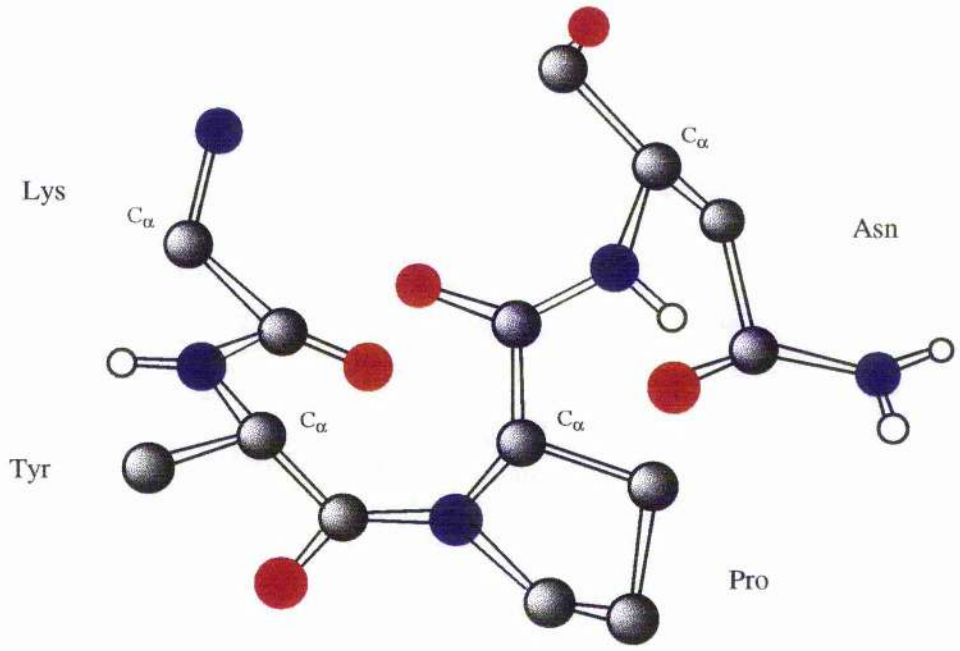
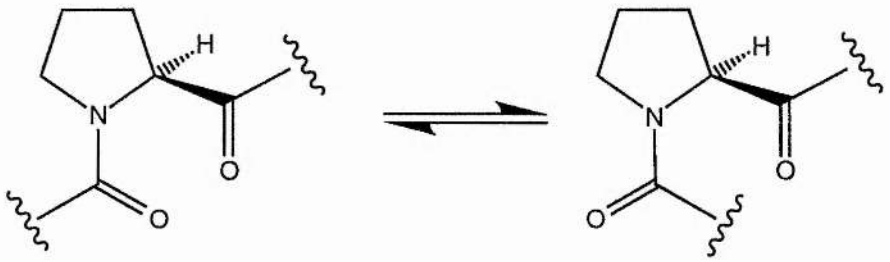


Figure 1.8: *A type VI β-turn*

1.3 Proline *cis/trans* isomerism

1.3.1 Introduction

In most peptide bonds, the side chain groups of the amino acids are held as far apart as possible in a *trans* configuration for the peptide bond. This is much preferable to the more crowded *cis* arrangement. For bonds involving proline, however, both *trans* and *cis* arrangements are possible (Scheme 1.4). The *trans* configuration is still slightly preferable to the *cis* form, the former being more energetically stable by approximately 10.5 kJ mol^{-1}.[81]



trans *cis*

Scheme 1.4: *X-Pro peptide bond isomerism*

Proline differs from the other amino acids in that it contains a secondary rather than a primary amine group. Also, the side chain is bonded to both the amine nitogen and the α-carbon, resulting in a cyclic structure as shown above. This has several structural implications for peptides and proteins containing proline residues. Since there is no amide hydrogen, hydrogen bonding to other residues is impossible and the five-membered pyrrolidine ring also imposes rigid constraints on the N-C^{α} rotation. Uniquely, proline residues have a relatively high intrinsic probability (0.1 to 0.3) of having the *cis* rather than the *trans* conformation of the preceding peptide bond as compared with less than 0.001 for other amino acids.[82] The activation energy barrier for *cis-trans* isomerization is also less for proline: 55 kJ mol^{-1} compared with 85 kJ mol^{-1} at other peptide bonds.[83] This is partly due to the greater length of the X-Pro bond (1.36Å instead of 1.33Å),[84] which results from the redistribution of charge and lack of resonance stabilization caused by the loss of the amide hydrogen. The amino acid residue preceding proline must also be given special consideration because the bulky pyrrolidine ring restricts the available conformational space.

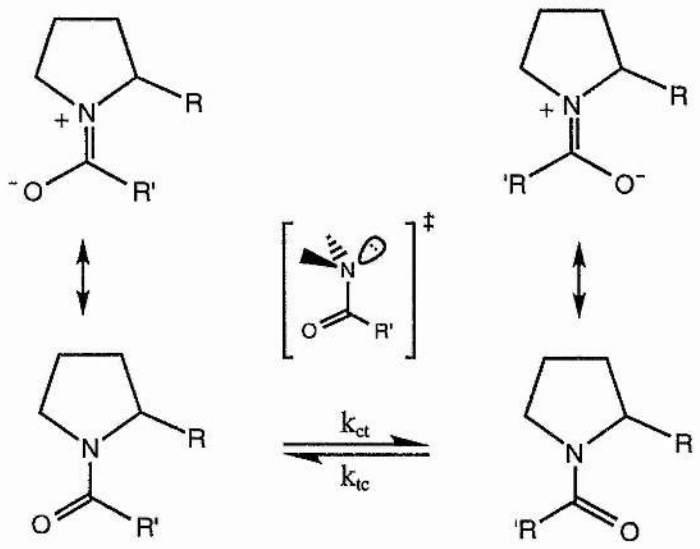
Proline residues are, therefore, recognized as being of special significance in their effect on chain conformation and the process of protein folding.[85]

It is clear that the presence of proline in a polypeptide chain imposes restrictions on its susceptibility to most peptidases.[86,87] These restrictions are related both to the presence of an X-Pro bond and the geometry of this bond. The unique effect of proline in a polypeptide chain suggests that this residue functions in a regulatory capacity to protect the peptide and perhaps to initiate the processing of biologically active peptides.[88]

NMR studies have shown that two distinct folded conformations of staphylococcal nuclease coexist in solution and that these two states can interconvert directly without passing through an unfolded state.[89] It has been proposed that *cis/trans* isomerism at a single peptide bond between a proline and its preceding residue might be the origin of the conformational multiplicity. Proline 117 was identified as a likely candidate for the site concerned. In the crystal structure of the nuclease-inhibitor complex, proline 117 is *cis* and is located in a β-reverse turn.[90] A mutant protein, in which proline 117 was replaced by glycine, was constructed in order to test this proposal and lent powerful support to the theory due to the fact that alternative conformations are not observed in the spectrum of this mutant. Similarly, two-dimensional ^{1}H NMR spectroscopy has shown that the calcium-binding protein, calbindin D_{9k} has two native conformations in solution.[91] This is also believed to be the result of the *cis/trans* isomerism of proline 43, one of its four proline residues.

1.3.2 Mechanism of amide bond rotation

The amide bond is one of the most important molecular structures and a proper representation of its chemical characteristics is essential to explain the chemical and biochemical properties of peptides and proteins. The nature and reactivity of the amide bond has been mainly explained by the concept of chemical resonance which suggests a partial π character for the C-N bond, due to a transfer of electrons from the nitrogen to the carbonyl group.[92,93] This hypothesis explains a large number of molecular properties of amides, such as their planarity, IR C=O frequencies, and the short length of the N-C bond.



cis isomer *trans* isomer

Scheme 1.5: *Mechanism for proline* cis/trans *interconversion*

From a chemical point of view, the *cis/trans* interconversion involves the disruption of the amide resonance and a high rotational barrier would be expected due to the loss of the π conjugated bond (Scheme 1.5). Recently however, the classical model for amides has been questioned due to molecular mechanics studies which have pointed out some inconsistencies.

1. In dialkyl substituted amides, the nitrogen atom deviates significantly from the plane of the three carbon atoms to which it is attached.[94]
2. The electronic population at the carbonyl oxygen does not noticeably change when the amide bond rotates.[93] The force constants for the C-O bond do not change either and only small changes in the C-O distance occur on rotation.

These findings would suggest that the role of the oxygen is not as important as is suggested by the classical mechanism and that although the rotation of amides is a highly unfavourable process, the energy barrier is perhaps not as great as is generally assumed.

1.3.3 Factors affecting the *cis/trans* ratio

It has been shown from the calculation of activation energy parameters that the barrier to isomerization is almost entirely enthalpic in all solvents studied.[95] There are several major factors which affect the kinetics of *cis/trans* interconversion in small peptides and also disturb the equilibrium by preferential stabilization of one of the two conformers.

1. **Solvent effects**

   Although it has been reported that the *cis/trans* rotational barrier is very similar for small peptides in a range of solvents,[96] more recent studies have shown that the rate constants for both *cis* to *trans* and *trans* to *cis* interconversions are significantly larger in non-polar, aprotic solvents.[97-101] This observation is consistent with a mechanism involving a less polar transition state as shown in Scheme 1.5. The fact that protic solvents restrict isomerization suggests that the barrier for interconversion is proportional to the strength of hydrogen bonds formed to the amide oxygen.[98] This may be of biological relevance, since it has been suggested that the enzymic acceleration of peptidyl prolyl isomerization (section 1.4) stems largely from the hydrophobic environment of the active site.[102] It has been noted for two nonapeptides that the rate of *cis* to *trans* isomerization is accelerated to a greater extent on going from water to methanol than the rate of *trans* to *cis*.[100] It seems that changes in conformation take place which destabilize the *cis* isomers in the less polar solvent. Other studies on di- and tripeptides also observed an increase in the relative amount of the *cis* isomer present going from less polar to more polar solvents.[97,99,101] These studies have suggested that polar, protic solvents tend to stabilize the *cis* conformation probably due to hydrogen bonding. Although the *trans* isomer has a generally more extended conformation, there is lower solvent accessibility to the amide carbonyl in this isomer relative to the *cis* form.

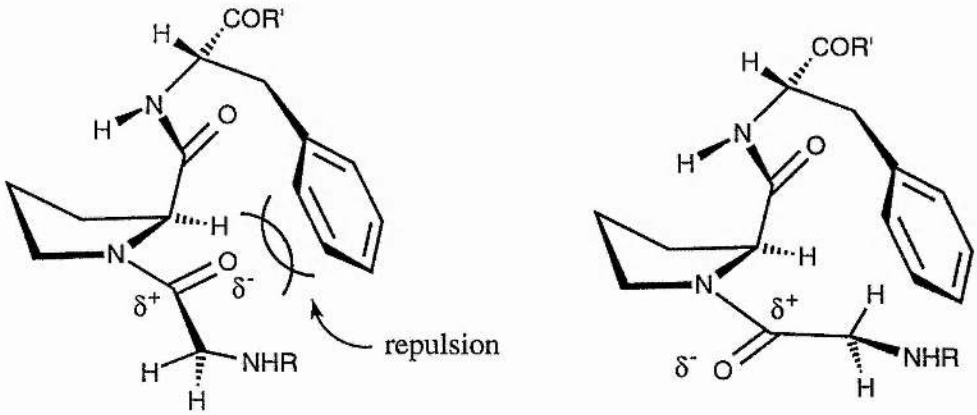
2. **pH effects**

   In addition, the *cis/trans* ratio in small peptides has been shown to be extremely sensitive to pH.[97,99,101] The ionized forms of various small peptides have shown strong preferences for one particular rotational isomer. In basic solution, it has been observed that peptides which have proline as the carboxyl terminal residue exhibit a large increase in the proportion of the *cis* isomer. This is most readily explained on the basis of electrostatic repulsions between the deprotonated proline carboxylate oxygen and the preceding amide carbonyl

oxygen. The two carbonyls are closer in the *trans* conformation leading to greater destabilization. The proportion of the peptides in the *trans* form increases at low pH which would seem to be the result of favourable hydrogen bonding between the protonated acid group and the neighbouring amide carbonyl.

3. **Stereoelectronic effects**

The nonapeptide, bradykinin (RPPGFSPFR) contains the sequence Ser-Pro-Phe and the Ser-Pro bond is known to exist in both the *cis* and *trans* isomeric forms where the *cis* conformation is populated by about 10% of the total ensemble. Substitution of a glycine residue for serine-6 was shown by NMR studies to lead to a dramatic increase in the population of the *cis* isomer to 35%.[103] Studies with model tripeptides have demonstrated that the phenylalanine residue adjacent to the proline is also important in determining the *cis/trans* ratio.[104] A stereoelectronic argument has been suggested to explain the unusually high *cis/trans* ratio exhibited by peptides containing the Gly-Pro-Phe sequence. It has been proposed that the *trans* conformation is destabilized due to a repulsive interaction between the glycine carbonyl oxygen and the π electrons in the phenylalanine aromatic ring (Fig. 1.9a).[105] This destabilizing interaction could not exist in the case of the *cis* isomer (Fig. 1.9b). It should be noted that Gly-Pro-Phe is the sequence of the FMDV polyprotein at the 2A/2B junction.



(a) *trans* isomer (b) *cis* isomer

Figure 1.9: *Electrostatic interactions in Gly-Pro-Phe*

4. **Artificial receptors**

The *cis/trans* ratio can be perturbed by the addition of artificial receptors which preferentially bind to one of the two rotamers. The terephthaloyl receptor shown below binds tightly to the appropriately spaced carboxylic acid groups of succinyl proline which shifts the equilibrium towards the less favourable *cis* isomer (Fig. 1.10).[106]

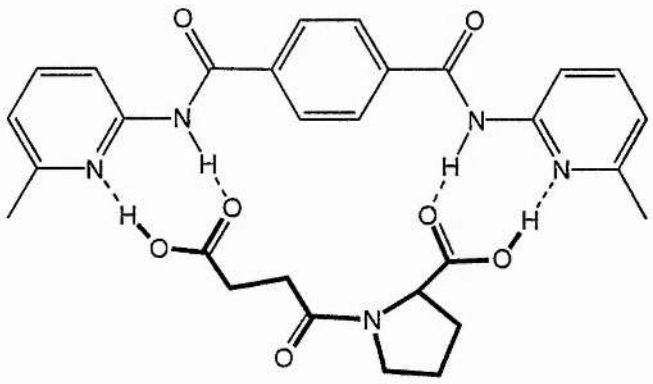


Figure 1.10: *Prolyl peptide coordinated to artificial receptor*

1.3.4 Examples of non-prolyl *cis* peptide bonds

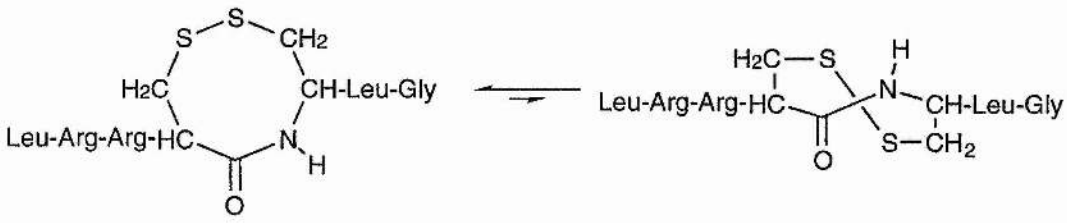
In those instances where the *cis* isomer is present, a proline residue is almost invariably a participant in the peptide linkage.[107] There are however, several structural and environmental conditions which can cause the occurrence of a *cis* peptide linkage between amino acids other than proline.

1. N-methyl amino acids which occur naturally in various plant metabolites display a strong preference for the β VI folded conformation containing a *cis* amide bond.[108] As well as providing increased resistance against biodegradation and increased hydrophobicity, the N-methylation of peptides induces various structural perturbations. One of these is the stabilization of the *cis* conformation due to the presence of a tertiary amide bond.[109]

2. *cis* Amide bonds have been observed within strained cyclic tetrapeptides which do not contain proline or any other N-substituted amino acid residues.[110] In these cases, the conformation is largely governed by the restriction of ring closure. In a study of a weakly constrained cyclic hexapeptide (*cyclo*-AFVKWF), the presence of a small amount (6%) of a second isomer was

detected.[111] This was shown to be a conformer containing a *cis* peptide bond between lysine and tryptophan, thought to arise due to steric interactions with the bulky side chain of an adjacent valine residue.

3. An example of a molecular constraint that generates a *cis* peptide bond is the presence of two adjacent cysteine residues which are connected *via* both amide and disulfide linkages.[112] The eight-membered ring resulting from the formation of the disulfide bond compels the peptide linkage to adopt the *cis* arrangement since the *trans* conformation is highly strained as shown below (Scheme 1.6).



cis *trans*

Scheme 1.6: *Isomerism of Cys-Cys peptide bond*

4. Peptide-membrane interactions can be energetically more important than the steric constraints usually associated with the *cis* arrangement. When in the presence of membrane-mimicking SDS micelles, a *cis* peptide bond was observed between isoleucine and lysine in Bombolitin I, a linear 17-residue peptide (IKITTMLAKLGKVLAHV), isolated from bumble bee venom.[113] This is the first and only report of a *cis* amide bond in a completely linear peptide without proline or N-substituted amino acids and is thought to arise from the association of the peptide with the hydrophobic surface of the micelles. In aqueous solution, the peptide adopts a random conformation but when SDS is added above the critical micelle concentration, the peptide adopts a well-defined α-helix. This study clearly indicates that the membrane can have a dramatic effect on the conformation of a peptide and underlines the importance of examining peptides within an environment similar to the conditions wherein they exhibit activity.

1.4 Prolyl *cis/trans* isomerases

1.4.1 Introduction

The *cis-trans* isomerization of X-Pro peptide bonds is thought to be involved in the refolding reactions of many proteins.[82] These processes are slow and they frequently determine the overall rate of folding, particularly for small, monomeric proteins.

In a study by Schmid, it was observed that various unfolded proteins consist of a mixture of conformational species that differ vastly in their rates of folding.[114] Brandts *et al.* first put forward a hypothesis which provided a plausible molecular explanation for this phenomenon.[82] They suggested that the fast-folding (U_F) and slow-folding (U_S) species differed in the geometry of one or more X-Pro peptide bonds. Normally, in the native protein, N, each prolyl peptide bond has a unique conformation, either *cis* or *trans*. However, during unfolding, ($N \rightarrow U_F$), the conformational restraints of the native state vanish and these bonds become free to isomerize slowly in the $U_F \rightarrow U_S$ reactions.

$$\mathbf{N} \underset{\text{folding transition}}{\rightleftharpoons} \mathbf{U_F} \underset{\text{prolyl isomerisation}}{\rightleftharpoons} \mathbf{U_S}$$

The relative concentrations of U_F and U_S at equilibrium are determined by the number of prolyl peptide bonds, their isomeric state in the folded protein and by the *cis/trans* equilibria in the denatured protein.

1.4.2 Immunophilins

The search for an enzymic activity that would catalyse prolyl peptide bond isomerization began soon after the proposal that prolyl isomerism was responsible for different folding rates in proteins. Success came in 1984 when Fisher *et al.* isolated an enzyme from porcine kidney that catalyzed the *cis-trans* isomerization of prolyl bonds in oligopeptides.[115] The enzyme promoted a 180° rotation about the X-Pro amide bond without cleavage or formation of a covalent bond and was termed peptidyl-prolyl *cis-trans* isomerase (PPIase), rotamase or conformase. In addition to

this activity towards small substrates, PPIase was also found to accelerate slow steps in the folding of several proteins.[116,117]

A surprising result emerged from the sequencing of PPIase.[118] It was found to be identical with cyclophilin, the major high-affinity binding protein in the cell for the immunosuppressive drug cyclosporin A.

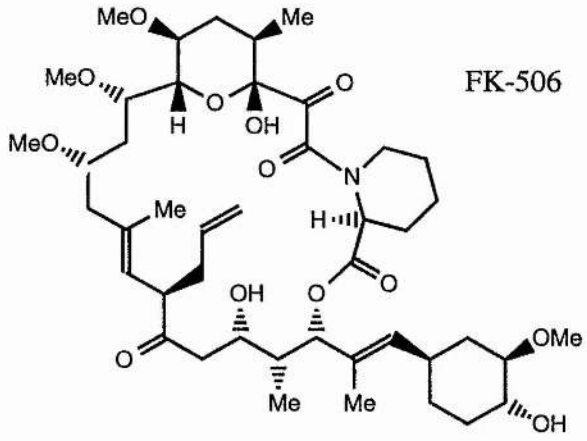
After the initial discovery of cyclophilin by virtue of its isolation from porcine kidney,[115] additional members of the cyclophilin family were detected in many cells and tissues.[119-121] All members of the cyclophilin group possess prolyl isomerase activity and are able to accelerate the slow steps in the folding[122] and unfolding[123] of various proteins. The PPIases act as true catalysts and, unlike molecular chaperones, they do not depend on the presence of stoichiometric amounts of the helper protein or low molecular weight additives.[124] PPIases were also the first enzymes known to be specialized for the catalysis of conformational changes.

A second class of prolyl *cis-trans* isomerases, the FK506-binding proteins (FKBPs), was discovered in 1989.[125-127] These proteins are inhibited by the immunosuppressants FK506 and rapamycin, but they do not bind cyclosporin A.[128] Schreiber *et al.* introduced the term "immunophilins" for the family of PPIases in order to point out their role as receptor proteins in the mammalian immune system.[129]

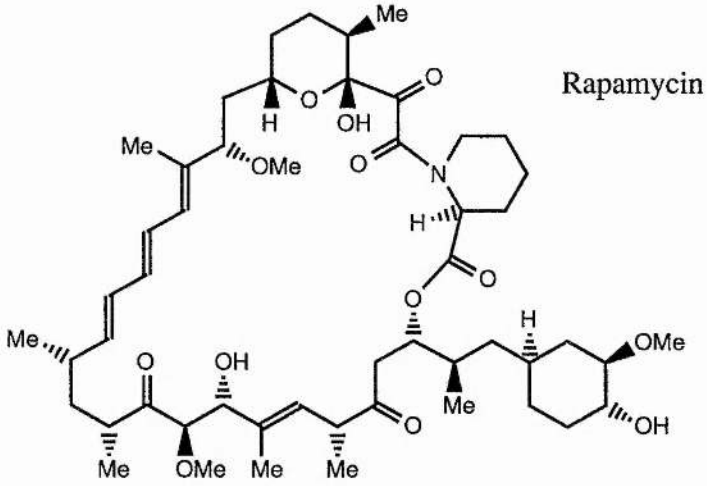
Immunophilins occur in all organisms, both prokaryotic and eukaryotic and in various subcellular departments, including the endoplasmic reticulum and the periplasm of *E. coli.*[130] Although the exact cellular concentrations of the immunophilins are not known, they are abundant proteins and cyclophilins are estimated to constitute up to around 0.4% of the total cellular protein.[125]

1.4.3 Immunosuppressants

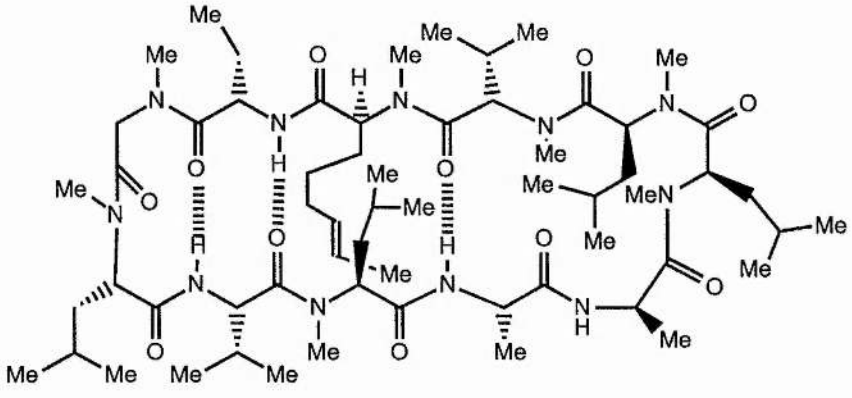
In surgical transplant procedures, immunosuppressive agents are used to control rejection of the new tissue. Advances in the area of organ transplants and bone marrow therapy have been greatly aided by the use of such immunosuppressants as cyclosporin A and the more recently discovered FK506 and rapamycin (Fig. 1.11). FK506 and rapamycin have been the subjects of considerable synthetic[131-134] and degradative[135,136] studies.



FK-506



Rapamycin



Cyclosporin A

Figure 1.11: *Immunosuppressant drugs*

These life saving drugs appear to interact with receptor proteins that are located in the cytoplasm, such as cyclophilin and FK506-binding protein. As both FKBP and cyclophilin have PPIase activity, the medical relevance of the immunophilins has been a driving force for new insights into proline *cis-trans* isomerism.

Cyclosporin A, a cyclic undecapeptide, has found widespread clinical use in the prevention of graft rejection following bone marrow and organ transplantation.[137] FK506 is a structurally novel macrolide of bacterial origin and, in addition to its immunosuppressive activity, it also has significant *in vivo* neuro- and nephrotoxicity. It has been demonstrated to have potent immunosuppressive activity at concentrations several hundredfold lower than cyclosporin A.[126] Rapamycin inhibits T-cell activation at concentrations comparable to those of the structurally related FK506, yet with mechanisms that are strikingly different from those mediated by FK506, and thus cyclosporin A.[129] The mechanisms of immunosuppression mediated by FK506 and cyclosporin A appear to be remarkably similar, suggesting that these unrelated structures act on a common receptor or on similar molecular targets.

The immunosuppressive activity of these drugs involves repression of a discrete set of lymphokine genes. This acts on the immune system by inhibiting initial steps involved in the activation of T-cells, thymus-dependent lymphocytes that act directly to incapacitate or destroy tissues bearing foreign antigens or invading microorganisms.

Both FKBPs and cyclophilins catalyze the interconversion of the *cis* and *trans* isomers of prolyl residues in peptides, and the appropriate immunosuppressants inhibit this activity. FK506 and rapamycin are inhibitors of FKBP but not of cyclophilin, likewise, cyclosporin A does not inhibit FKBP, yet is a potent inhibitor of cyclophilin. The FKBPs and cyclophilins are members of an emerging class of proteins which regulate T-cell activation and other metabolic processes, perhaps by the recognition and isomerization of proline-containing epitopes in target proteins.

1.4.4 Biological function of prolyl *cis/trans* isomerases

Defining the biological role of the immunophilins and their substrates may illuminate the process of protein trafficking in cells and the mechanisms of signal transmission through the cytoplasm.

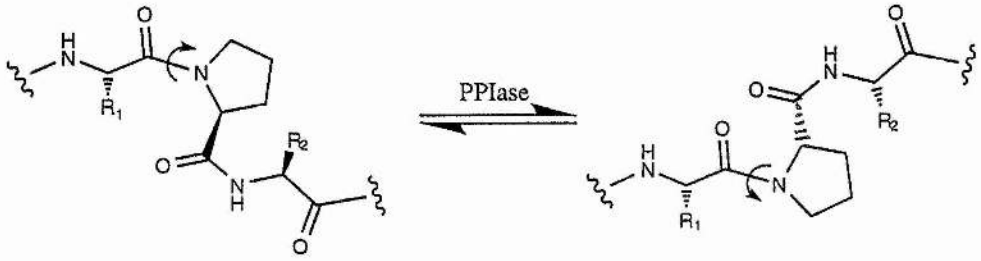
The immunosuppressive drugs, cyclosporin A, FK506 and rapamycin are highly effective inhibitors for PPIases. Surprisingly, these compounds effect the signal cascade of T-cells but not through enzyme inhibition; the inhibitor-enzyme complexes themselves are the active agents.[138] In the case of cyclosporin A and FK506, the complex of drug and immunophilin binds to calcineurin, a calcium-activated phosphatase that is essential for T-cell signalling.[139,140] Rapamycin is thought to behave differently, hampering T-cell activation at a later stage in the immunoresponse.[141]

Intensive efforts are being made to understand the role of prolyl *cis/trans* isomerases and of their complexes with immunosuppressants. Three-dimensional structures of the cyclophilin-cyclosporin A and FKBP-FK506 complexes have been determined by X-ray crystallography[142,143] and NMR spectrometry.[144,145]

At present, it is too early to speculate whether the primary cellular function of prolyl *cis/trans* isomerases is correlated with processes that involve protein folding or with signal transduction pathways. Possibly, they are involved in all these processes and the isomerase activity is important for their regulation. Whether such a multitude of tasks is accomplished by the same enzymes or rather by specialized prolyl isomerases is not known.

1.4.5 Mechanism of prolyl *cis/trans* isomerases

The reaction catalyzed by prolyl *cis/trans* isomerases differs from most other enzymic reactions. It is not associated with the formation or breakage of covalent bonds, intermediates do not accumulate and no cofactors are known to participate in catalysis (Scheme 1.7). Therefore, none of the usual experimental methods are available to help probe the mechanism.[146] On the other hand, prolyl isomerization provides the enzymologist with a rare opportunity since both the catalyzed and uncatalyzed reactions occur at measurable rates.



Scheme 1.7: *Peptide bond rotation catalyzed by prolyl isomerases*

The main barrier to prolyl *cis/trans* isomerization is the resonance stabilization energy possessed by the C-N imide bond. The task of a prolyl isomerase is, therefore, to develop an enzymic strategy that will result in the lowering of this barrier. There are two general mechanisms which have been suggested: nucleophilic catalysis and catalysis by distortion. The first of these two mechanisms is covalent catalysis in which a nucleophile attacks the carbonyl carbon to form a tetrahedral intermediate (Fig. 1.12a). In this intermediate, resonance stabilization of the C-N bond has been removed and collapse of the intermediate can produce either the *cis* or *trans* X-Pro bond. The alternative non-covalent catalysis mechanism is one in which the enzyme utilizes its binding energy to stabilize selectively the transition state with a 90° rotation of the amide bond out of planarity (Fig. 1.12b). This mechanism is dependent on the binding energy between enzyme and substrate, since the substrate is considerably distorted and strained.

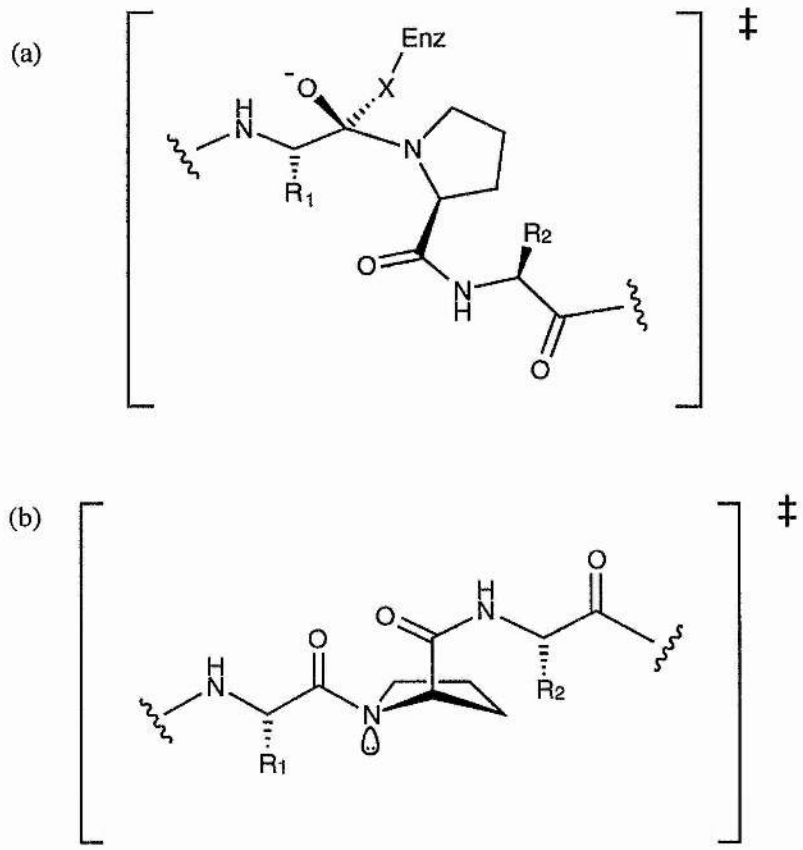


Figure 1.12: *Transition state structures for (a) covalent catalysis mechanism and (b) mechanism of catalysis by distortion*

Initial studies on cyclophilin led to the suggestion that an enzymic thiol group was important for catalysis and a nucleophilic mechanism was proposed involving a covalent intermediate, possibly a hemiortho-thioamide.[147] Site-directed mutagenesis of human cyclophilin allowed the systematic replacement of all four cysteine residues with alanine. Since all four mutant enzymes were fully active in rotamase and binding assays, cysteine was ruled out as a participating residue in catalysis.[148]

Additional mechanistic studies with both cyclophilin[127] and FKBP[149] strongly support a mechanism involving catalysis by distortion in which the PPIase binds and stabilizes a transition state that is characterized by partial rotation about the C-N amide bond. This mechanism is supported by crystal structures of FK506 and cyclosporin A that reveal substructures which mimic the twisted amide bond of the suggested substrate-enzyme complex.[129,149]

The energy that is required to distort this bond out of planarity with the C=O bond, thereby destroying the resonance stabilization of the amide linkage, is supplied by favourable transition state binding interactions between enzyme and substrate. The structural basis for the ability of the prolyl isomerases to stabilize this transition state must await further analyses. The bond distortion need not be completely the result of mechanical deformation but could also be due to desolvation or electrostatic destabilization.[146]

The energetic cost for bond distortion is reflected in the large entropy of activation ($-T\Delta S^{\ddagger}$ = 56 kJ mol^{-1} at 300K) that is required to distort and constrain the substrate to a conformation that allows optimal interaction with the enzyme. Once the entropic bill has been paid, the reaction occurs with little enthalpic cost ($\Delta H^{\ddagger}$ = 16.6 kJ mol^{-1}).[150]

This mechanism is unique in enzymology. PPIases appear to use none of the usual catalytic "devices" that enzymes typically employ (*e.g.* general-acid, general-base or nucleophilic catalysis) but, rather, utilize the intrinsic binding energy that is available from substrate-enzyme interactions in stabilizing the transition state for catalysis.

1.5 Peptidomimetics

1.5.1 Introduction

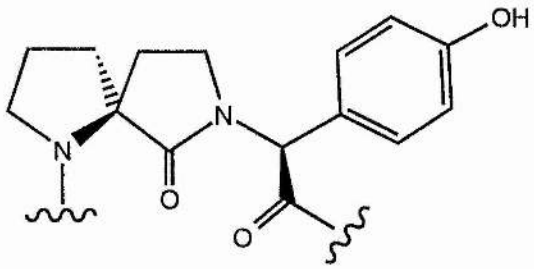
Peptides influence a multitude of physiological processes and their role in the appearance or maintenance of various diseases has been unequivocally proven. Agents that can imitate or block the biological function of bioactive peptides can be considered as aids for the investigation of the system and also as therapeutic agents. It has been shown that due to their pharmacological properties, such as degradation by peptidases or poor bioavailability, peptides could be employed as drugs in only a few cases. To solve this problem, compounds that act as substitutes for peptides in their interaction with receptors, peptidomimetics, have been synthesized.[151]

In the many studies aiming to obtain specific and superactive peptide analogues, attempts have been made to design molecules which are able to favour a particular conformation by introducing geometrical constraints. As well as cyclization and the insertion of nonproteinogenic amino acid residues, the alteration of the peptide bond can also influence conformational properties and biological activity. These modified molecules are known as peptidomimetics or pseudopeptide analogues.[152-154]

Previous investigators have sought to design and incorporate constraints into bioactive molecules which mimic features of peptidyl secondary structure while retaining or enhancing biological potency and if possible, introducing desirable properties such as enhanced oral absorption or increasing *in vivo* half life. A secondary structure mimetic is a building block that forces a defined secondary structure after its incorporation into a peptide.

1.5.2 β-turn mimics

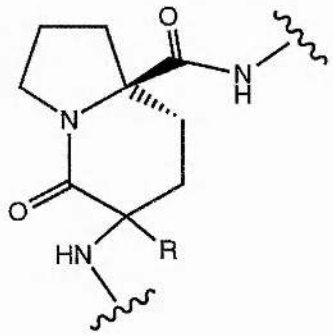
The β-turn is a structural motif common in many biologically active cyclic peptides[78,80] and it has been postulated that the activity of many linear peptides depends on the existence of a β-turn.[155,156] For this reason, there has been great interest in the synthesis of β-turn mimetics but the majority of these have been found to be inactive after incorporation into model peptides. The γ-spirolactam segment (**2**) has been used to replace Pro-Tyr units as a type II β-turn mimic.[157]



(**2**): *Spirolactam β-turn mimetic*

Although computational studies show that distances and torsion angles for peptides containing the spirocyclic unit are in good agreement with those expected for type II β-turns, the results of biological tests have proved disappointing and often the peptide analogues fail to bind to recognition sites.[158]

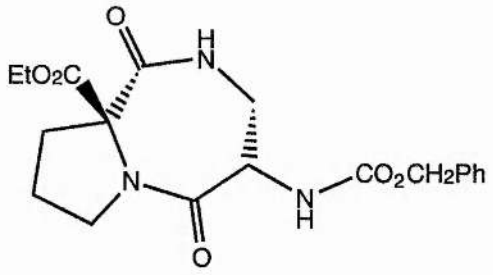
The type VI β-turn contains a *cis* X-Pro peptide bond and various *cis* amide bond surrogates have been designed to mimic this feature. The *cis* Gly-Pro mimetic (**3**) has been synthesized and incorporated into model peptides using solid-phase methods.[159] The bicyclic system allows retention of the peptide backbone and amino acid side chains in positions closely similar to those likely to occur in native *cis* X-Pro type VI turn conformations. When tested, however, analogue peptides containing the unit proved less active than the natural peptides.



(**3**): cis *Gly-Pro mimetic unit*

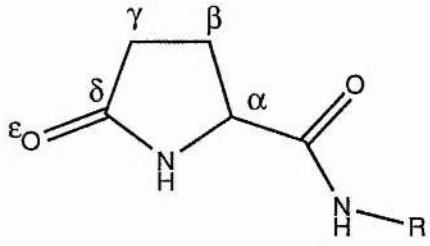
A very similar type VI β-turn mimic, (**4**), has recently been synthesized by Curran and McEnaney.[160] In this case, the two adjacent α carbons are tethered by a lactam methylene (CO-NH-CH_2) linker which provides additional constraint to the bicyclic framework compared with the alkyl (CH_2CH_2) linker used in the previous

example. This bicyclic unit has not yet been incorporated into larger peptides, so its biological activity as a β-turn mimetic is unknown.



(**4**): cis *X-Pro dipeptide*

Paul *et al.* have shown using molecular dynamics calculations that the pyroglutamic acid residue (Glp), (**5**), is a natural motif for a *cis* peptide bond.[161] An effective type VI β-turn mimetic can be made by starting from a Glp residue and extending it appropriately from its C^{γ} atom. The Glp residue contains a *cis* amide bond in the ring - the $C^{\gamma}C^{\delta}O^{\varepsilon}NHC^{\alpha}$ chain can be visualized as a peptide unit $C^{\alpha}CONHC^{\alpha}$ with the side chains at both C^{α} atoms being fused at C^{β}.



(**5**): *Pyroglutamic acid residue*

Modelling studies on tripeptides containing the Glp residue show that one of the minimum energy structures obtained is a type VI β-turn. Experimental work on these systems has yet to be carried out.

1.5.3 *Cis* and *trans* X-Pro mimics

Various conformationally constrained proline analogues have been prepared which favour the *cis* conformation.[162-165] Compounds such as *anti*-5-methylproline (**6**), 5,5-dimethylproline (**7**) and 2,2-dimethylthiazolidine carboxylic acid (**8**) have been shown to exhibit a relatively high percentage of the *cis* isomer in model compounds. Such modified proline surrogates should prove useful in assessing the role of *cis* X-Pro peptide bonds in biologically active peptides.

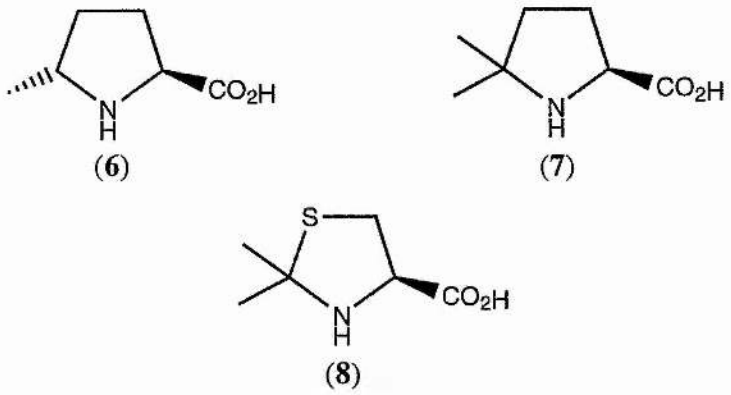
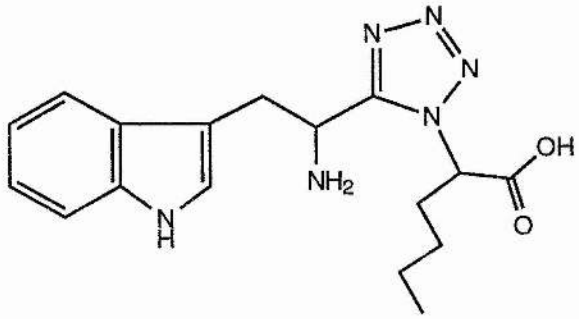


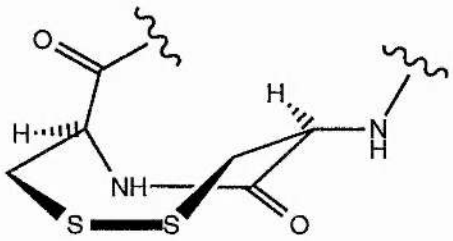
Figure 1.13: *Proline analogues which stabilize the* cis *X-Pro conformation*

Unlike its *trans*-olefinic counterpart which can be considered as an ideal mimic for the *trans* peptide bond, the synthesis of *cis* olefins as *cis* amide bond mimics is not synthetically attractive because of facile isomerization into the stable *trans* form. 1,2-disubstituted pyrroles[165] and 1,5-disubstituted tetrazole units[167] have been used as *cis* amide bond surrogates. Compound (**9**) was prepared as a *cis* Trp-Nleu mimetic but showed less activity than the parent peptide, suggesting that other structural characteristics are important for biological activity.[168]



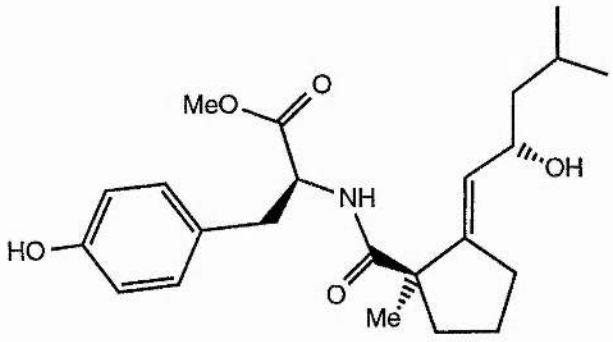
(**9**): cis *Trp-NLeu mimetic*

A novel disulfide constrained cyclic segment, (**10**), is formed by the closure of two neighbouring cysteinyl residues and was incorporated into a cyclic hexapeptide (*cyclo*-PFWKTF) to replace a Phe-Pro unit.[169] NMR studies have established that in solution the amide bond of this unit was essentially all *cis*. The amide bond in the mimetic compound is fixed in the *cis* geometry by the eight-membered ring, with apparently little disturbance of stereoelectronic parameters. Biological assays of the artificial peptide showed retention of high potency which would seem to be confirmation of the unit as a good mimetic for *cis* amide bonds.



(**10**): cis *Phe-Pro mimetic*

trans Prolyl peptidomimetics have also been prepared. These possess a *trans* substituted alkene in place of the proline peptide bond. The Leu-Pro-Tyr analogue (**11**), was synthesized by Andres *et al.* and proved to be a potent inhibitor of the prolyl isomerase activity of FKBP 12.[170] This shows that the simple carbocyclic system possesses considerable potential in mimicking the prolylamide substructure of polypeptide species. It is recognised by the enzyme and binds tightly to it but since the amide bond is removed, so is the possibility of peptidyl isomerization.



(**11**): trans *X-Pro peptide bond mimetic: Leu-Pro-Tyr analogue*

CHAPTER 2: RESULTS AND DISCUSSION

2.1 Synthesis and structural studies of 2A-containing oligopeptides.

2.1.1 Methods of structural analysis

Information on the three dimensional structure of peptides containing the FMDV 2A sequence may provide important insights into the relationship between structure and activity of this protein region. The solution phase conformations of small peptides (5-30 residues) are usually investigated by use of ^{1}H NMR spectroscopy[54,56,58] and CD (circular dichroism) spectropolarimetry.[51,171,172]

CD is the differential absorption of circularly polarized radiation by a non racemic sample.[173] In peptides, it is the interaction between adjacent amide groups during light absorption which accounts for most of the CD bands. The key structural elements that are responsible for the wavelengths and intensities of the bands are the relative orientation of the neighbouring amide groups and the distances between them. When the peptide molecules organize themselves into a helical shape and start to form intramolecular hydrogen bonds, absorption of radiation by the numerous properly oriented chromophores is enhanced. This leads to the appearance of more intense CD bands at characteristic wavelengths. CD spectroscopy allows the presence of α-helices to be inferred with greater certainty than that of any other structural feature.[59]

Two-dimensional ^{1}H NMR spectroscopy provides an important tool for identifying preferred conformers in otherwise unstructured peptides.[174] This stems from its usefulness in detecting highly localized conformations that are not easily observable by other methods, and from its ability to identify the particular residues involved in a given structural motif. NMR can be used to discern conformations that resemble regular secondary structures of proteins, such as helices[47,56,57] and tight turns.[60,64,171] The NMR approach to studies of peptides is largely oriented around the ability of nuclear Overhauser enhancement (nOe) experiments to provide data on internuclear distances.[55] 2D NOESY spectra are used to identify protons physically close in space within large molecules which tumble slowly enough to lie in the negative nOe region.[175] In characterizing conformations, however, it must be appreciated that a peptide is frequently in dynamic equilibrium between a large number of conformers. Therefore, at any one time the peptide population in solution contains multiple conformers, only some of which are revealed by NMR.

2.1.2 Influence of different solvents on peptide structure

Small, biologically active, linear peptides often have significant secondary structure, usually as turns or helices, in organic solvents while they are largely unstructured in aqueous solution.[46,49,50,53,61] The influence of organic solvents on the experimentally determined secondary structure of a peptide is thought to be consistent with a role for the membrane-lipid interface in stabilizing an active conformation.[62] Although it is not clear why such peptides acquire a specific conformation at the lipid interface, reduction in the dielectric constant and the activity of water molecules is obviously part of the driving force.

2,2,2-Trifluoroethanol (TFE) has been widely used as a structure-inducing cosolvent. As well as enhancing the solution structure of small protein fragments, TFE can induce the formation of stable conformations in peptides which are otherwise randomly arranged in aqueous solution.[46,50,51] Several properties have been suggested to be responsible for the observed secondary structure stabilization.

- The dielectric constant of TFE is about one third that of water and more closely resembles conditions in the interior of a protein.[49] Thus, one would expect strengthened interactions between charged groups in this medium.

- Differences in solvent acidity and basicity between TFE and water are thought to change the relative stability of hydrogen bonds.[46] The hydrogen bonding of amide protons to the solvent is decreased in TFE, which strengthens intramolecular hydrogen bonds and therefore stabilizes secondary structure.

- TFE is a less polar or more hydrophobic solvent than water. It interrupts hydrophobic interactions and can act as a denaturant of tertiary and quaternary structures.[46]

TFE is not a helix-inducing solvent in the sense that it will induce helix formation independently of the sequence. It is rather a helix-enhancing cosolvent, which stabilizes helices in regions with some helical propensity. This can be fairly accurately determined by secondary structure predictions.[62]

2.1.3 Peptide synthesis and structural analysis

In order to investigate the mechanism of the self-cleavage of the 2A region of the FMDV polyprotein, we decided to synthesize model peptide fragments. Molecular biological studies carried out by Ryan *et al.* were used to decide the lengths of peptides to be prepared.[67] In these studies, recombinant DNA experiments were carried out in which specific portions of the FMDV polyprotein were deleted in order to investigate the region involved in the 2A/2B cleavage. Initial results indicated that a truncated version of the 2A region in which the first 6 residues were deleted, exhibited self cleavage.[67]

1D→ 2A ← 2B
-LLNFDLLKLAGDVESNPGPFF-

The first oligopeptide prepared, ALAGDVESNPGPF (**12**), was thirteen amino acids in length and included the ten C-terminal residues of the 2A region. The peptide was prepared by solid-phase synthesis[176-178] (see Appendix 1) using Fmoc protection.[179]

The peptide (**12**) was analyzed by HPLC using a reversed phase C_{18} column and was found to elute as a single peak (>90%) using an H_2O/ acetonitrile gradient and 214 nm detection. The FAB mass spectrum of (**12**) gave the expected molecular ion ($[M + H]^+ = 1273$) using a glycerol matrix. The ^{13}C NMR spectrum of (**12**) was recorded in ^2H_2O and the carbonyl region was found to contain the expected 16 peaks (Fig. 2.1).

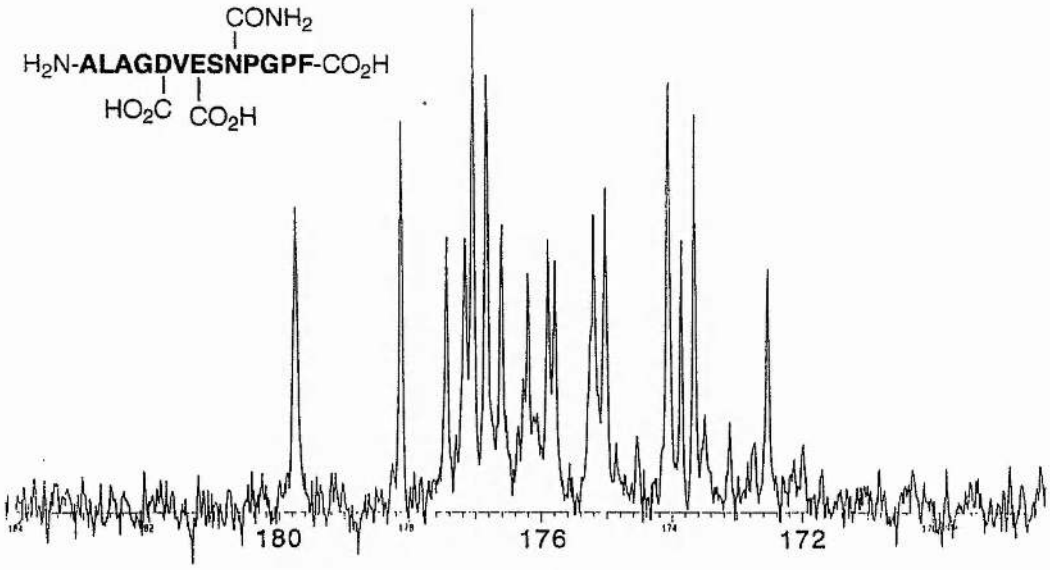


Figure 2.1: *Carbonyl region of the ^{13}C NMR spectrum of (12)*

The synthetic peptide (**12**) was found to be insoluble in DMSO and TFE. A range of 2D NMR experiments were therefore performed in water in order to see if the peptide contained any secondary structural elements in this medium.

From the NMR results, it seemed that (**12**) was predominantly in an unfolded conformation in H_2O/ ^2H_2O (9:1) (see Appendix 2).

- The proton chemical shifts in the ^1H NMR spectrum corresponded very closely to the known random coil shifts for amino acid residues in water.[174]

- The ROESY spectrum contained a series of sequential $\alpha(i)$-NH(i+1) crosspeaks from Ala^1 to Val^6, but there were no NH(i)-NH(i+1) or $\alpha(i)$-$\alpha(i$+3) interactions which are characteristic of a helical conformation.[55,175]

- Strong nOe crosspeaks were observed between the α-protons of Asn^9 and Gly^{11} and the δ-methylene protons of Pro^{10} and Pro^{12} respectively. This is conclusive evidence that both X-Pro amide bonds are predominantly in the *trans* conformation.

Using the same synthetic methodology (see Appendix 1), a 23 residue peptide, NFDLLKLAGDVESNPGPFFFSDF (**13**), was prepared. This peptide contained the entire 2A sequence as well as seven residues from the adjacent 2B region. The reason for the inclusion of the 2B portion was twofold. First, it was thought that an extended C-terminal might be important for cleavage activity and second, it was hoped that the additional hydrophobic groups might make the peptide more soluble in organic solvents. It is known that short linear peptides are more likely to exhibit structure in organic solvents than in aqueous solution.[63,64]

The peptide (**13**) was analyzed by HPLC using a reversed phase C_{18} column and was found to elute as a single peak (>90%) using an H_2O/ acetonitrile gradient (Fig. 2.2). The FAB mass spectrum of (**13**) gave the expected molecular ion ($[M + H]^+ = 2577$) using a NOBA matrix.

Peptide (**13**) was found to be water soluble and, unlike (**12**), when the pH was adjusted to 7.8, it also became sparingly soluble in both DMSO and TFE. We therefore decided to investigate the structure of (**13**) in different solvents, using both CD and NMR techniques (see Appendix 2).

The CD spectra of (**13**) in H_2O (Fig. 2.3a) and H_2O/ TFE (1:1) (Fig. 2.3b) were recorded at pH 7.8 and 35 °C. On changing the solvent from H_2O to the H_2O/ TFE mixture, a large maximum band appears at 190 nm, the minimum absorption peak ($\pi\pi^*$ transition) shifts from 200 to 205 nm and the intensity of the negative ellipticity at 222 nm ($n\pi^*$ transition) increases.

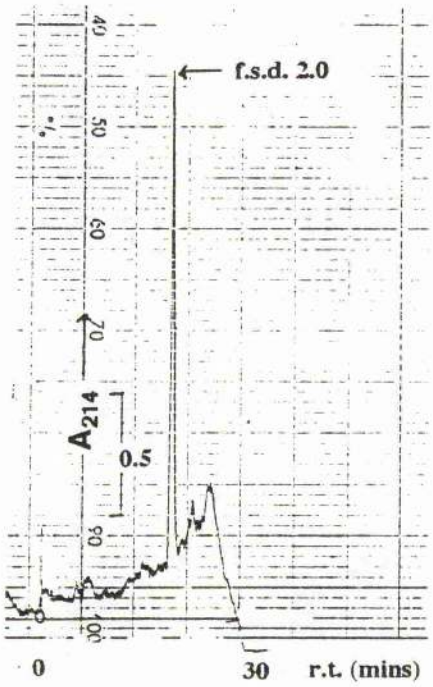


Figure 2.2: *RPHPLC trace of (13)*
(A_{214} - Absorbance at 214 nm, f.s.d. - full scale deflection, r.t. - retention time)

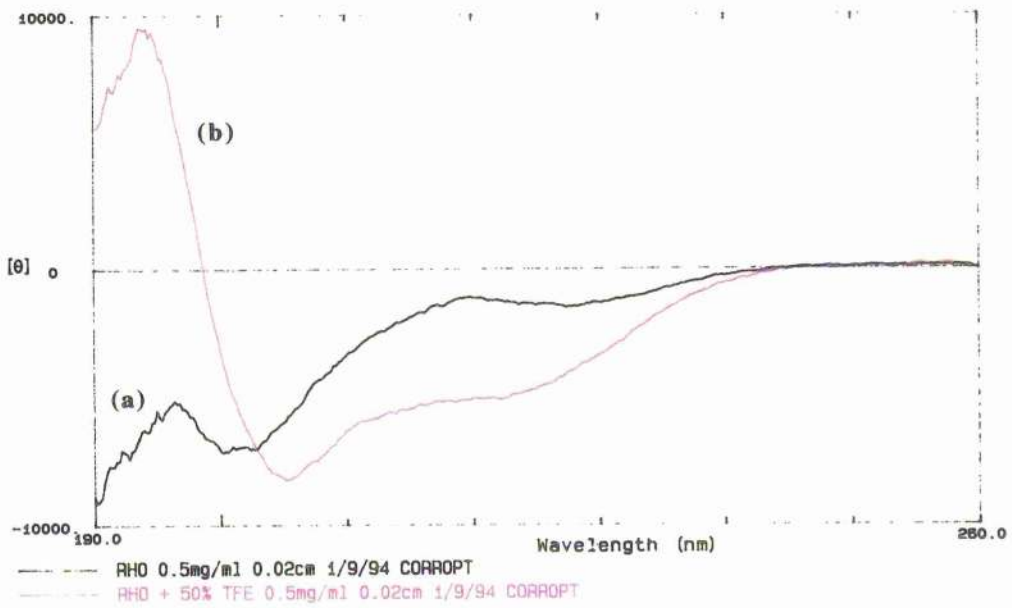


Figure 2.3: *CD spectra of (13) in (a) H_2O and (b) H_2O/ TFE (1:1)*

The shape of the red curve (Fig. 2.3b) is characteristic of a peptide containing α-helical secondary structure.[48,173] The large intensity of the minimum at 205 nm relative to the minimum at 222 nm implies that the conformation is a mixture of α-helical and random coil structures.

The amount of helical structure at 37 °C was estimated by two methods. First, the procedure of Provencher and Glöckner was used.[59] This is a computational method which is based on describing the spectrum as a combination of the spectra of various proteins of known secondary structure. This method gave a value of 10% for the amount of α-helix present. An alternative approach is to use the value of the molar ellipticity at 208 nm, $[\theta]_{208}$ (measured in units of degrees cm^2 $dmol^{-1}$) to calculate the amount of α-helical structure present in the peptide. Using a simple formula,[49] a value of 11.2% for the estimated α-helical content was obtained.

Although CD spectra can reveal overall characteristics of the secondary structure, no local information is available on the conformation of peptides. Therefore, analysis of 2D NMR spectra is essential to identify the secondary structure of each residue and to obtain a clearer picture of the conformation of the peptides.

High resolution NMR methods made it possible to investigate specific residues in (**13**) and a range of 2D ^{1}H NMR experiments were run in H_2O/ ^2H_2O (9:1), d_6-DMSO and d_2-TFE.

The first step in structure determination was to assign each proton in the molecule to a specific resonance in the NMR spectrum. 2D TOCSY experiments were employed to connect *via* scalar coupling all resonances belonging to different protons in the same amino acid residue.[174] Although the TOCSY spectrum enables assignment of each resonance to a type of amino acid based on the observed coupling patterns, it does not give the sequential location of the amino acid. The sequential asssignments were obtained from observation of inter-residue interactions between protons in the NOESY spectrum, *e.g.* NH(i)-NH(i+1) or α(i)-NH(i+1).[55]

A 2D NOESY spectrum yields a map of protons that are spatially close irrespective of bonding networks. Different elements of regular secondary structure have different anticipated patterns of short distances. For instance, in a regular α-helix, the distance between sequential amide protons is quite short (d_{NN} 2.8 Å), whereas this distance is considerably longer in a β-sheet (d_{NN} 4.3 Å) or extended structure (d_{NN} 4.2 Å).[174] The α-helix is also characterized by close approach between residues i and (i+3), and between residues i and (i+4) which result in medium range nOe interactions. These characteristic interactions, NH(i)-NH(i+1) and α(i)-NH(i+3) can be used to determine regions which have helical as opposed to extended conformations.

The NMR spectra of (**13**) in H_2O/ ^2H_2O (9:1) showed the peptide to be unstructured in that medium. The proton chemical shifts closely corresponded to known random coil values[174] and no NH(i)-NH(i+1) crosspeaks were observed in the NOESY spectrum.

In the NOESY spectra of (**13**) in both DMSO and TFE, however, seven sequential NH(i)-NH(i+1) interactions were observed from Leu^4 to Val^{11}, suggesting an area of secondary structure such as an α-helix (Fig. 2.4). No α(i)-NH(i+3) interactions were detected which may indicate that the helix is not very stable. The number of NOE interactions, especially α(i)-NH(i+3), is indicative of the size of the helical population and the stability of the helix. The amide region of the TOCSY spectrum of (**13**) is

shown for comparison (Fig. 2.5). This contains crosspeaks arising from interactions between the side chain amide protons of the asparagine residues and the ε-amine group of the lysine residue.

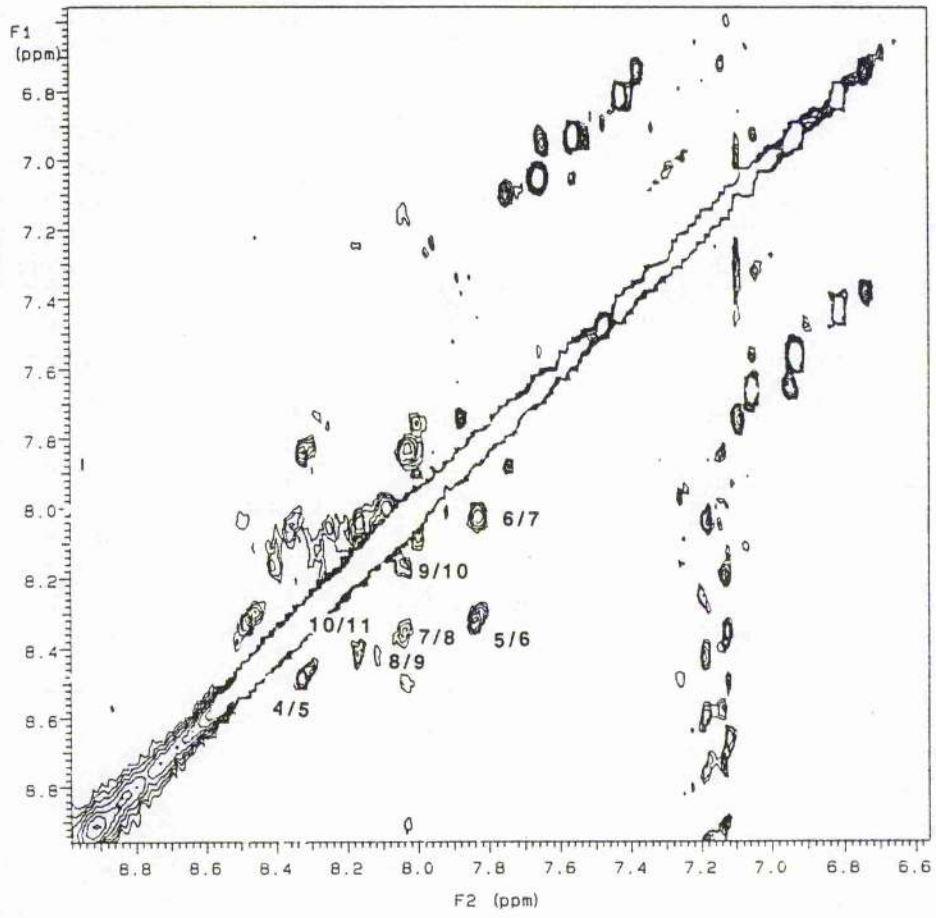


Figure 2.4: *Amide region of NOESY spectrum of (13) in TFE*

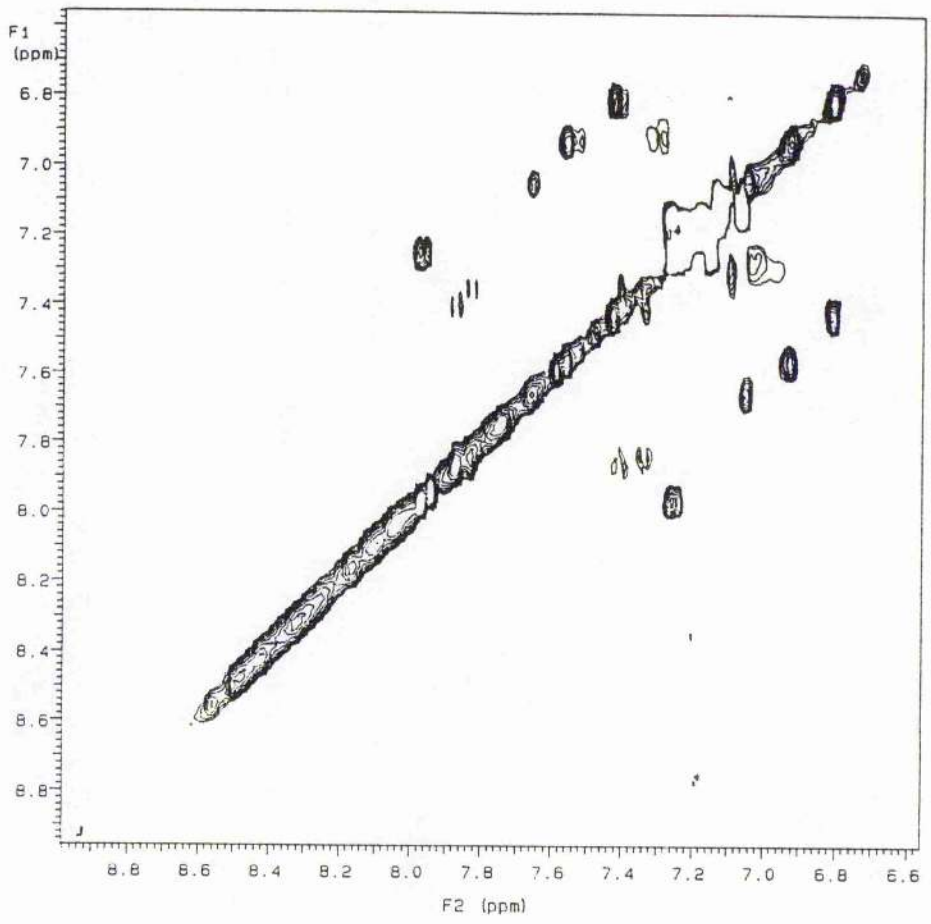


Figure 2.5: *Amide region of TOCSY spectrum of (**13**) in TFE*

2.1.4 Conclusions

The results reported here present experimental evidence that the conformation of peptide (**13**), which contains the complete 2A sequence is solvent dependent, changing from a random coil in H_2O to a partial α-helix in TFE or DMSO. The helical area is believed to be located in the N-terminal region of the peptide. These results are in agreement with molecular modelling studies which indicated that the 2A sequence might form an α-helix of seven to eleven residues followed by a reverse turn in close proximity to the scissile bond (see Fig. 1.5, p. 23).[36,67]

The adoption of some form of structure upon exposure to a more hydrophobic environment (such as DMSO or TFE) is typical behaviour for peptides which are active at membrane surfaces.[62] In the case of the FMDV polyprotein, the 2A region appears to display proteolytic activity when still closely associated with host cell ribosomes (see section 1.1.8, p. 15). Ribosomes are large globular protein complexes and may provide an apolar or interfacial environment, similar to those of membrane surfaces.

Molecular modelling studies indicated that the presence of a *cis* X-Pro bond would cause a sharp turn in the 2A peptide backbone which would enable the scissile Gly-Pro bond to be positioned close to potentially nucleophilic residues (see section 1.2.2, p. 21).[67] The occurence of *cis/trans* isomerism about an X-Pro bond in a peptide or protein is often manifested as a set of minor peaks present in the 2D NMR spectra.[63,91,180,181] Diagnostic evidence for the presence of a *cis* peptide bond is the observance of an $\alpha(i)$-$\alpha(i+1)$ NOESY crosspeak.[174] No such evidence of conformational multiplicity was apparent in the NMR spectra of (**12**) or (**13**), and it can be assumed that the two X-Pro bonds, present in both synthetic peptides exist in the more usual *trans* configuration.

2.2 Synthesis and studies on short 2A peptide fragments

2.2.1 NPGP - Introduction

Since 1990, there have been various reports in the literature that the synthetic tetrapeptide NPGP (Asn-Pro-Gly-Pro) spontaneously cleaves when incubated in alkaline buffer.[2,14,16,43] The NPGP sequence corresponds to the 2A/2B cleavage site of the FMDV polyprotein and these accounts suggest that the inherent instability of the NPGP fragment may be a driving force in the scission of the viral protein.

These literature reports consist only of brief comments in papers and reviews on the general topic of viral proteases and no experimental details have yet been published. We decided to synthesize several protected and deprotected forms of the tetrapeptide fragment in order to investigate its alleged autocatalytic properties and to examine the accuracy of these literature reports. Several points about the purported cleavage have been noted in these short accounts.

- The autocatalytic reaction is reported to be carried out most efficiently in slightly basic reaction mixtures (*i.e.* pH 8.5), as might be expected for an authentic physiological event.[16]

- Thus far, the spontaneous degradation has reportedly been observed only for peptides with unmasked amino-termini.[43]

- The cleavage of the tetrapeptide is thought to produce two dipeptide fragments: NP and GP.[14] In contrast, the cleavage in the viral polyprotein occurs at the NPG/P junction, suggesting a different mode of cleavage.

- Palmenberg *et al.* have reported that a synthetic peptide of the sequence NPGP-NH_2 (Asn-Pro-Gly-Pro-amide) spontaneously hydrolysed to yield predominantly NP and GP-NH_2 fragments.[43]

In contradiction to the last point,[43] the present study has shown conclusively that the dipeptide amide GP-NH_2 is unstable, particularly in a basic environment and rapidly reacts to form a diketopiperazine (see Section 2.4, p. 80).

2.2.2 Solution phase synthesis of the NPGP tetrapeptide

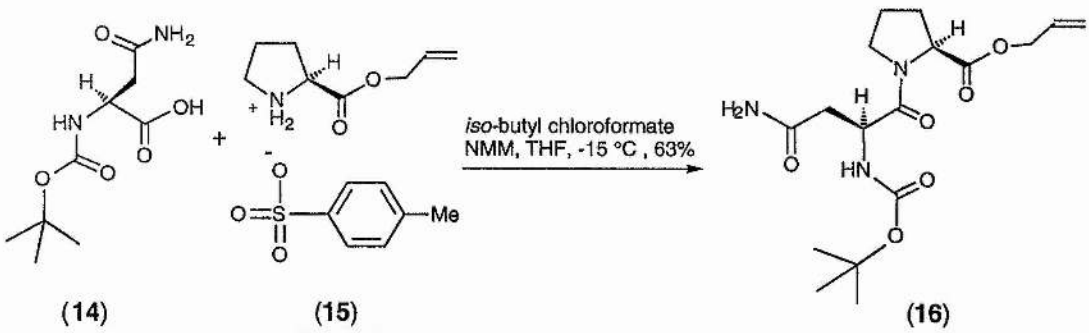
The tetrapeptide (NPGP) was prepared using classical solution phase peptide chemistry. A convergent approach was employed, involving the coupling of two dipeptide fragments: NP and GP (Scheme 2.2).

There are numerous excellent methods for synthesizing peptides in the solution phase, including the use of acid chlorides or carbodiimides for peptide couplings.[182] We synthesized peptide fragments using the mixed anhydride method of peptide coupling. This method was chosen for its advantages in speed, yield and relative purity of the products. The mixed anhydride method has been studied in detail to increase yield and minimize the possibility of racemization.[183,184] The reaction uses alkyl chloroformates as peptide forming reagents. A suitably protected amino acid reacts with the alkyl chloroformate in the presence of a base, to form an anhydride. A second suitably protected amino acid is then added which reacts with the anhydride to form the peptide linkage.

An important consideration is that of racemization, which can be minimized depending on the steric bulk of the base and the choice of solvent. It has been found that methylamine reacts the fastest, but causes maximum racemization.[183] Base, solvent, temperature and alkyl chloroformate variations have been examined, and it was found that the best results are obtained using a combination of NMM, *iso*-butyl chloroformate and dry THF.[184] This combination of reagents was used for the synthesis of the NPGP tetrapeptide where possible.

Reactions involving proline and asparagine are known to be problematic due to the hindered nature of the proline amine group and the poor solubility of asparagine compounds. The Asn-Pro dipeptide proved to be difficult to synthesize in good yield largely due to these factors and various reaction conditions and solvents were tried out in order to improve the yield.

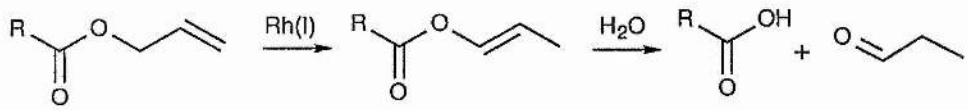
(2*S*)-Proline methyl ester hydrochloride (**21**) was found to be insoluble in THF and could not be successfully coupled to ^{t}BOC-(2*S*)-asparagine (**14**) for this reason. In contrast, (2*S*)-proline allyl ester *p*-toluene sulfonate (**15**) was found to be soluble in THF and was reacted successfully with ^{t}BOC-(2*S*)-asparagine (**14**) using the mixed anhydride methodology to give ^{t}BOC-Asn-Pro-allyl ester (**16**) in 63% yield after recrystallization.



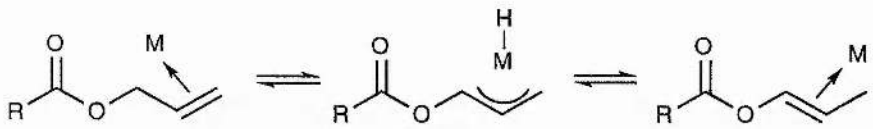
(14) (15) (16)

Scheme 2.1: *Synthesis of ^{t}BOC-Asn-Pro-allyl ester (16)*

Deprotection of the allyl ester using conventional Wilkinson's catalyst[185] or $Pd(PPh_3)_4$,[186] proved unsuccessful, as was the use of lithium dimethylcuprate.[187] Rhodium(I) complexes such as Wilkinson's catalyst are known to catalyze the isomerization of the allyl moiety as shown below which leads to the liberation of the carboxy function when reacted with water.



It is generally accepted that migration of hydrogen occurs with involvement of a π-allyl complex as shown below.



It seems likely that the allyl group of the protected dipeptide (**16**), was not sufficiently accessible to the bulky metal reagents for coordination to occur. This may be due to intramolecular hydrogen bonding involving the asparagine side chain as has been shown for several small peptides containing the Asn-Pro pair.[188,189] This restriction of conformational flexibility would greatly affect the reactivity of the allyl ester group (Fig. 2.6).

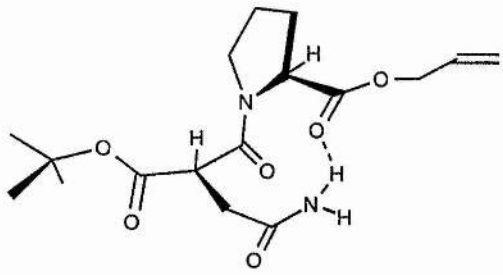
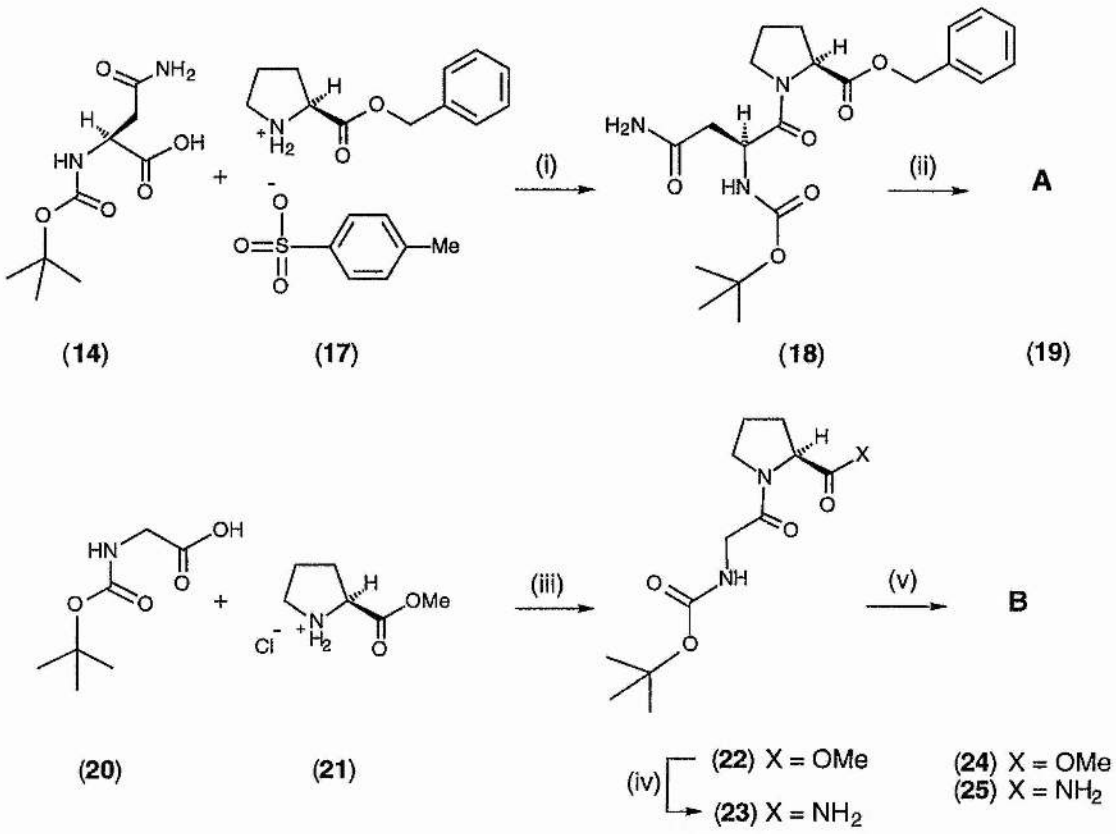


Figure 2.6: *Proposed conformation of ᵗBOC-Asn-Pro-allyl ester (16)*

As an alternative, (2*S*)-proline benzyl ester *p*-toluene sulfonate (**17**) was reacted with ᵗBOC-(2*S*)-asparagine (**14**) in an analogous manner to that above to give ᵗBOC-Asn-Pro-benzyl ester (**18**) in 79% yield after recrystallization, m.p. 111-13 °C (lit.,[190] 115-16 °C) (Scheme 2.2). The benzyl ester was then removed by catalytic hydrogenation to give the deprotected dipeptide (**19**) in 87% yield, m.p. 110-15 °C (lit.,[190] 108-16 °C). Both ^{1}H and ^{13}C NMR spectra of (**19**) showed the complete disappearance of aromatic signals.

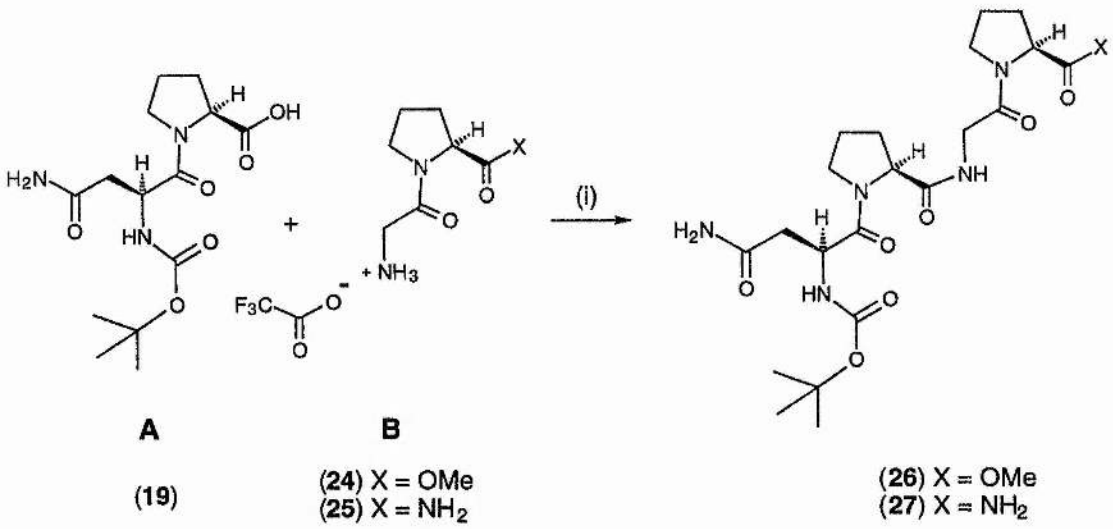
The second dipeptide, ᵗBOC-Gly-Pro-OMe (**22**) was also prepared by the mixed anhydride coupling of (2*S*)-proline methyl ester hydrochloride (**21**) with ᵗBOC-glycine (**20**) in 89% yield. Removal of the ᵗBOC group was easily accomplished using trifluoroacetic acid to give the methyl ester (**24**) as a clear oil in quantitative yield.



Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF, -15 °C, 79%; (ii) H_2(g), Pd/C, MeOH, rt., 96%; (iii) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 89%; (iv) NH_3(g), MeOH, 72 h, 98%; (v) TFA, CH_2Cl_2, rt., 100%.

Scheme 2.2: *Synthesis of the dipeptide fragments (19), (24) and (25)*

As the mixed anhydride method was found to be unsuccessful for the final coupling step to produce the tetrapeptide, the coupling reagent PyBOP (benzotriazolyloxy-tris[pyrrolidino]-phosphonium hexafluorophosphate) was employed for the condensation of the two dipeptides, (**19**) and (**24**).[191] The tetrapeptide ^{t}BOC-Asn-Pro-Gly-Pro-OMe (**26**) was isolated as a white solid in 75% yield, m.p. 168-70 °C. The peptide was analyzed by HPLC using a reversed phase C_{18} column and was found to elute as a single peak (>98%) using an H_2O/ acetonitrile gradient and 214 nm detection.



Reagents and conditions: (i) PyBOP, NMM, THF/ DMF, rt., 50 min.

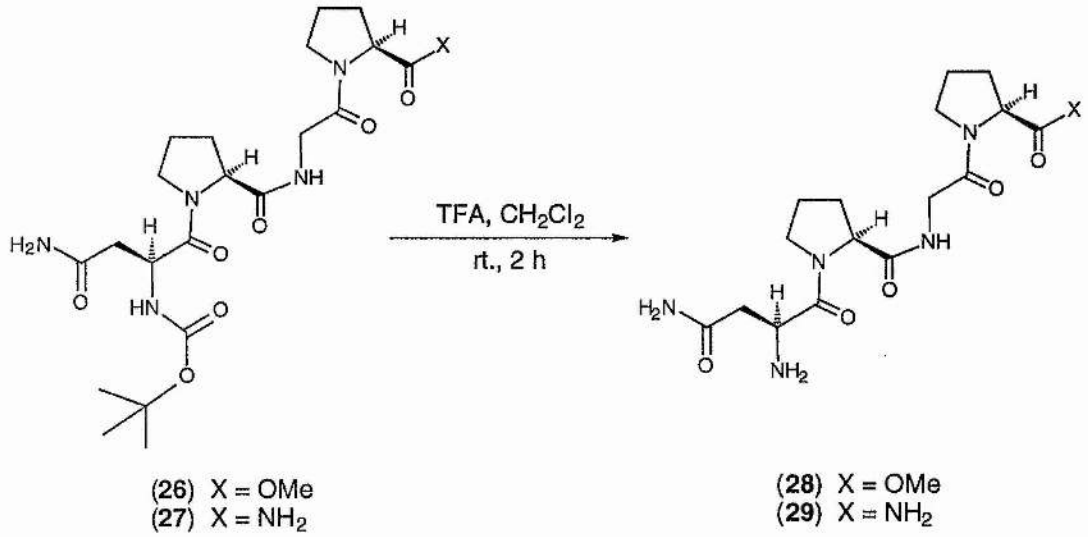
Scheme 2.3: *Synthesis of NPGP tetrapeptides (26) and (27)*

It is not clear which factors impeded the mixed anhydride coupling reaction between (**19**) and (**24**). If the proline carboxyl group of the NP dipeptide (**19**) is involved in an intramolecular hydrogen bond, this would render it rather hindered and unreactive. Phosphonium coupling reagents such as PyBOP[191] and the related PyBroP[192] have been shown to be extremely effective for the coupling of hindered peptides and, indeed, proved successful in this case.[193]

In order to ascertain the role (if any) of the terminal amino group in cleavage activity, we decided to prepare the peptide H_2N-Asn-Pro-Gly-Pro-OMe (**28**). After treatment of the tetrapeptide ester (**26**) with TFA, ^1H NMR showed from the disappearance of the large signal at 1.42 ppm that the ^{t}BOC protecting group had been completely removed. Since TFA is a strong acid, the initial product of the deprotection reaction was a TFA salt of the terminal amine but, the free amino peptide (**28**) was produced in basic solution.

In order to synthesize the corresponding tetrapeptide (**27**), the dipeptide ^{t}BOC-Gly-Pro-OMe (**22**) was reacted with a solution of methanol saturated with NH_3 gas (Scheme 2.2). The resulting dipeptide (**23**), m.p. 152-4 °C, was then deprotected and coupled with compound (**19**), as described above, to give the tetrapeptide ^{t}BOC-Asn-Pro-Gly-Pro-NH_2 (**27**). The amino terminus of this compound was also

deptrotected using TFA to give, in basic solution, the tetrapeptide H_2N-Asn-Pro-Gly-Pro-NH_2 (**29**). This was the exact form of NPGP prepared by Palmenberg *et al.*[43]



(**26**) X = OMe
(**27**) X = NH_2

(**28**) X = OMe
(**29**) X = NH_2

Scheme 2.4: *The NPGP tetrapeptides (**26-29**) examined for cleavage activity*

2.2.3 Investigation of structure and reactivity

The four tetrapeptides: (**26**), (**27**), (**28**) and (**29**) were incubated at pH 8.5 and 37 °C for 12 h, and no cleavage activity was observed, as determined using TLC and NMR spectroscopy. Even in the cases of (**28**) and (**29**) where the amino termini were deprotected, no spontaneous degradation was observed, in contrast to the report by Palmenberg *et al.*[43] Although the peptides showed no propensity to cleave, it was decided to examine the solution structure of the NPGP segment, given that it forms part of the sequence of the FMDV 2A region.

In order to find out if the NPGP fragment contained any interesting secondary structural features, the solution state conformation of the tetrapeptide (**26**) was studied in various solvents, using both NMR and CD. Some small proline-containing peptides have been shown to possess secondary structural features such as β bends. An example of this is Tuftsin, an immunoactive tetrapeptide (Thr-Lys-Pro-Arg) which has been shown by CD analysis to adopt a β-bend conformation in solution.[156]

The protected tripeptide Ac-Asn-Pro-Tyr-NHMe has also been shown to contain a β-bend which is thought to be stabilized by an intramolecular hydrogen bond between the asparagine side chain carbonyl and the terminal amide NH proton. The tripeptide ^{t}BOC-Asn-Pro-Ser-NHMe has also been shown to adopt an Asx (β III) type turn in both solution and solid phases which is not stabilized by an intramolecular hydrogen bond.[188] The tetrapeptide (**26**), however, was shown by CD spectra in both water and TFE to be in an extended conformation. There was no evidence of any secondary structure (Fig. 2.7).

The ^{1}H and ^{13}C NMR spectra of the tetrapeptide (**26**) revealed that a minor conformational isomer was present in 15% abundance. In the NOESY spectrum, the minor component gave rise to a cross peak between the glycine methylene protons and the α-proton of one of the proline residues (Fig. 2.8). This is diagnostic evidence for the presence of a *cis* amide bond.[174] The two proline residues were unambiguously assigned from the COSY spectrum and it was shown that the *cis* amide bond was between the glycine and the terminal proline residue.

The results from these studies did not provide any substantial evidence for the mechanism of cleavage of the 2A peptide. However, it was shown that the NPGP sequence has a reasonably high propensity for the *cis* conformation at the Gly-Pro, rather than the Asn-Pro, amide bond. This has important implications for the conformation of the larger 2A peptide which is cleaved at the Gly-Pro junction. A *cis* Gly-Pro bond would cause a reversal in the peptide chain direction and would enable the scissile bond to be situated close to potentially nucleophilic residues within the 2A segment (see Fig. 1.5, p. 23).

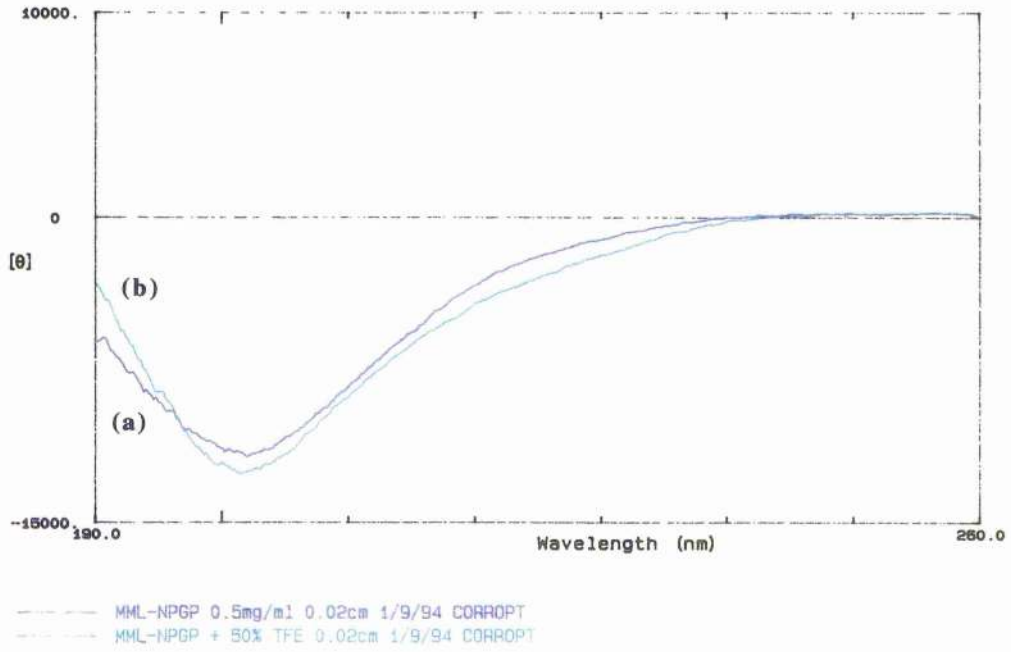


Figure 2.7: *CD spectra of (26) in (a) H_2O and (b) H_2O/ TFE (1:1)*

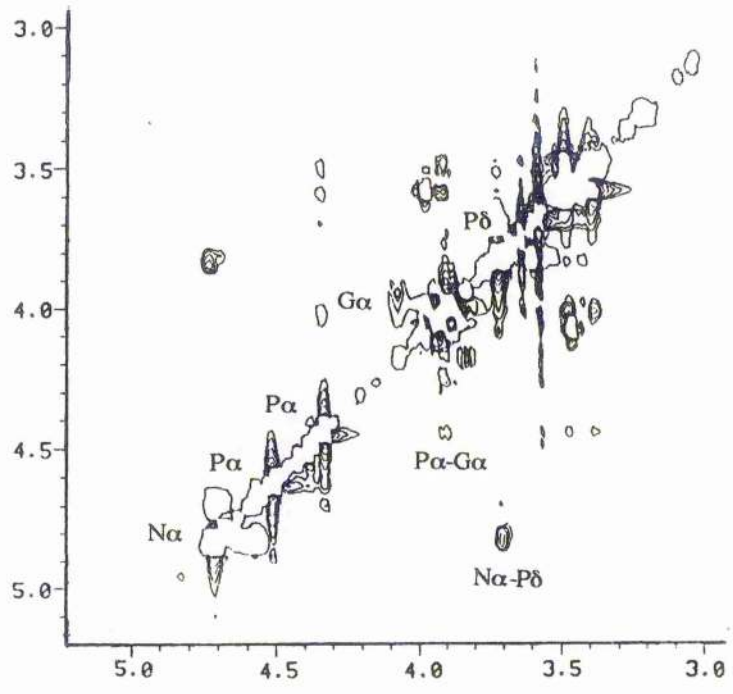


Figure 2.8: *ROESY spectrum of (26)*

2.2.4 GPF - Introduction

There have been numerous discussions in the literature on the factors which determine the *cis/trans* ratio of X-Pro peptide bonds. Previous studies have noted a correlation between the steric bulk in the side chain of the residues preceding or following proline and the *cis/trans* ratio.[83] From other results, however, it appears that the nature of the side chain in the amino acid preceding proline has a negligible effect on the *cis/trans* ratio.[104] Since simple steric arguments fail to explain the observed experimental results, there is a need for a more detailed understanding of this phenomenon.

London *et al.* have noted the unusually high *cis/trans* ratio exhibited by peptides containing the sequence Gly-Pro-Phe. They have proposed that the *trans* conformation is destabilized due to a repulsive interaction between the glycine carbonyl oxygen and the π electrons in the phenylalanine aromatic ring (see Fig. 1.9a, p. 33).[105] Since GPF is the sequence of the FMDV 2A polyprotein at the site of cleavage, it was decided to synthesize tripeptides containing this sequence to see if they did indeed possess large fractional concentrations of the *cis* isomer.

2.2.5 Synthesis of the GPF tripeptide

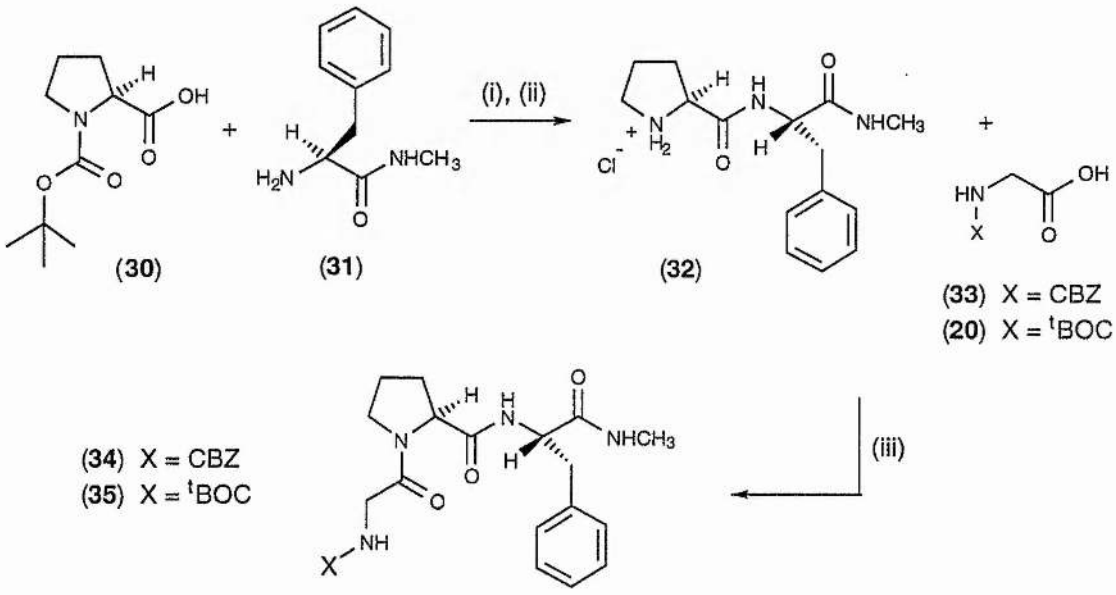
To avoid hydrogen bonding interactions involving terminal amine or acid groups, we opted to synthesize the peptides with protecting groups on both ends. Phenylalanine methylamide (**31**), was coupled to ^{t}BOC-proline (**30**), using the mixed anhydride method to give ^{t}BOC-prolyl phenylalanine methylamide. The ^{t}BOC protection was removed using HCl gas to give Pro-Phe-NHMe.HCl (**32**) in 94% after recrystallization, m.p. 212-14 °C. Subsequent coupling of the dipeptide (**32**) with CBZ-glycine (**33**), again using the mixed anhydride method, gave the tripeptide CBZ-Gly-Pro-Phe-NHMe (**34**) in 83% yield as a white solid, m.p. 62-6 °C.

In order to investigate the effect of different protecting groups on the *cis/trans* ratio, the dipeptide (**32**) was coupled to ^{t}BOC-glycine (**20**) as described above, to give ^{t}BOC-Gly-Pro-Phe-NHMe (**35**) in 82% yield, m.p. 75-7 °C.

Unfortunately, X-ray quality crystals of (**34**) and (**35**) could not be obtained so it was impossible to investigate the solid state conformations of these peptides.

From the ^{13}C NMR spectra of (**34**) and (**35**), it could be seen that the *cis/trans* ratios were not particularly high. In the case of (**34**), the *cis* isomer was present in approximately 7% abundance and this figure was slightly lower for (**35**) (5%). The *cis/trans* ratios for the GPF tripeptides were considerably lower than those of the NPGP tetrapeptide, (**26**) which was shown to contain around 15% of the *cis* isomer due to isomerization of the Gly-Pro bond.

Thus, it can be concluded that the presence of the phenylalanine residue does not have a significant effect on the *cis/trans* ratio of the GPF tripeptides, (**34**) and (**35**). From the studies on NPGP peptides, it seems that the residues directly preceding the Gly-Pro bond are more influential in stabilizing the *cis* conformation.



Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF, -15 °C; (ii) HCl(g), EtOAc, 0 °C, 20 min; (iii) NMM, *iso*-butyl chloroformate, THF, -15 °C.

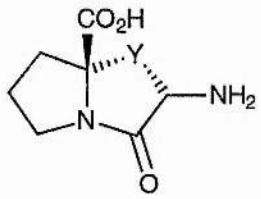
Scheme 2.5: *Synthesis of GPF tripeptides (34) and (35)*

2.3 Synthesis of *cis* prolyl peptides

2.3.1 Introduction

In order to verify the existence of a *cis* Gly-Pro bond in the 2A region of FMDV, we decided to prepare a variety of synthetic peptides which contained this structural element.

The obvious approach for holding an X-Pro amide bond in the *cis* configuration is to tether the α-carbons of the two amino acids with a linker Y (**36**).



(**36**)

Various bicyclic compounds with different linkers have been prepared and have been proposed as *cis* proline mimetics (see section 1.5, p. 44).[159,160,195] Although computational studies have shown that distances and torsion angles in these compounds are similar to those of naturally occurring *cis* X-Pro units, they have displayed little activity in biological tests.[159]

The main objective of this study was to synthesize a bicyclic compound where the linker Y could be broken under mild conditions to give the natural *cis* X-Pro dipeptide. This synthetic unit could then be incorporated into larger peptides and the linkage cleaved as the final step to release the free *cis* oligopeptide.

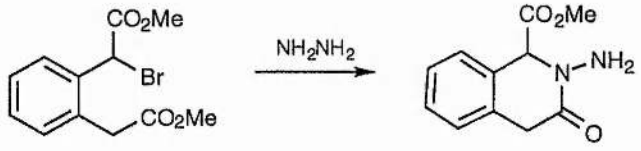
The thermodynamic parameters for amide bond rotation in a number of proline-containing peptides have been measured by a variety of methods.[96,109,150,196-200] The results all give values for the free energy of rotation about the X-Pro amide bond of 80-90 kJ mol^{-1}. Using the Eyring equation, it can be calculated that the isomerization is slow at room temperature with a half life of several seconds.[96] It is also known that the rate of *cis-trans* isomerization decreases with increasing length of the peptides on either side of the bond.[127] It was envisaged that the *cis* X-Pro unit could eventually be inserted into the larger 2A peptide and that the *cis* species could be retained long enough for self-cleavage to occur, presuming that is the criterion required.

2.3.2 Synthesis of a *cis* Gly-Pro mimetic

A bicyclic system with a five-membered pyrrolidine ring fused to a seven-membered triazepine ring in a [2,1-d] fashion was chosen as our initial target. Reductive cleavage of the larger ring across the nitrogen-nitrogen single bond would give an open chain *cis* peptide (Scheme 2.6).

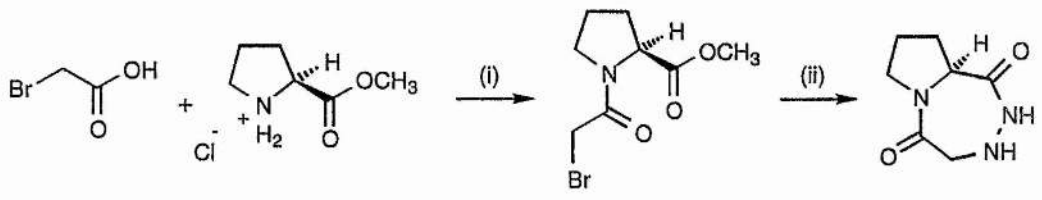
Scheme 2.6: *Synthesis of a* cis *X-Pro peptide from a bicyclic precursor*

It was planned to synthesize these systems from bifunctional proline dipeptides and suitably substituted hydrazines. This strategy would only be successful if the formation of seven, rather than six-membered rings was favoured. Previous studies by Cignarella *et al.*,[201,202] on the reactivity of simple hydrazines indicated a strong tendency for double alkylation at one nitrogen (Scheme 2.7).



Scheme 2.7: *Reaction of hydrazine with a bifunctional molecule*

Initially, the synthesis of a simple model compound was attempted in order to ascertain the feasibility of this approach. Bromoacetic acid, (**37**) was condensed with proline methyl ester hydrochloride (**21**), using mixed anhydride methodology to give compound (**38**) as a colourless oil. The reaction of (**38**) with excess hydrazine hydrate in refluxing ethanol was followed by TLC which showed the complete disappearance of starting material after 90 min. Upon cooling, a white precipitate formed, m.p. 265-6 °C. All spectroscopic and analytical data were consistent with the assignment of the product as the 5,7-bicyclic compound (**39**) (Scheme 2.8).



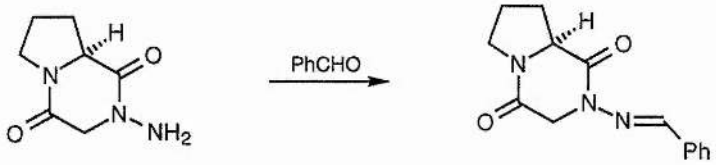
(37) (21) (38) (39)

Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 83%; (ii) $NH_2NH_2.H_2O$, EtOH, 78 °C, 1.5 h, 100%.

Scheme 2.8: *Synthesis of a bicyclic* cis *Gly-Pro mimetic*

It was not immediately obvious whether the structure of the product was the seven-membered triazepine (**39**) or the isomeric six-membered diketopiperazine, (**40**). Both compounds would give the same molecular ion in mass spectroscopy and their NMR spectra would be expected to be similar. The compound was eventually assigned as the seven-membered ring structure from its chemical reactivity and rather unusual NMR properties.

- The compound showed no reactivity with benzaldehyde. If a primary amino group was available, the isomer would give the benzylhydrazone derivative, (**41**), as shown below.[202]

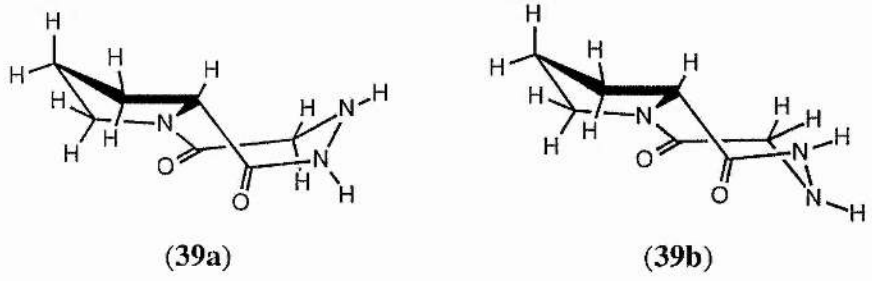


(40) (41)

- The ^1H NMR spectrum in D_2O showed only one set of signals but when the same sample was run in d_6-DMSO, two sets of signals of similar intensity were observed. There is a good deal of information available on the solution conformations of proline-containing diketopiperazines in various solvents and they are not known to display conformational isomerism in any solvent.[203-5]

- Two sets of signals of approximately equal intensity were also observed in the ^{13}C NMR spectra of (**39**) in both ^2H_2O and d_6-DMSO. Conformational isomerism has never been observed in the ^{13}C spectra of proline-containing diketopiperazines.[204]

The two sets of signals observed in the NMR spectra of (**39**) may be due to two interconverting conformational isomers which could arise from the geometry of the seven-membered ring. The two forms of (**39**) must be of similar stability with a reasonably high energy barrier of conversion between them. It has been calculated that the most stable conformations of cycloheptane rings are the twist chair and the twist boat which can only be interconverted through major deformations of bond angles.[206] Although the seven-membered ring system under discussion differs greatly from a cycloheptane molecule, molecular modelling studies indicate two energetically favourable conformations (Fig. 2.9). The isomer where the seven-membered ring is in a pseudo boat conformation (**39a**) was calculated to be slightly more stable (2.4 kJ mol^{-1}) than the pseudo chair conformation (**39b**). The ^1H NMR spectra appear to illustrate the different rates of interconversion in different solvents with only one set of signals observed in ^2H_2O and two distinct sets apparent in d_6-DMSO. The ring-flipping may be too fast in water to be discerned by NMR but is slowed down in DMSO to the extent that the different conformers are clearly discernable.



(**39a**) (**39b**)

Figure 2.9: *Two possible conformational isomers of (39)*

2.3.3 Synthesis of *cis* Ala-Pro mimetics

In order to extend the synthesis to the preparation of other *cis* X-Pro mimetics and to shed some light on the mechanism of formation of the seven-membered triazepine (**39**), we decided to synthesize the bicyclic *cis* Ala-Pro mimetics, (**49**) and (**50**) (Scheme 2. 9).

(2*S*)-Bromopropionic acid (**44**) was prepared by the method of Freudenberg from (2*S*)-alanine (**42**).[207] Reaction of the α-halo acid with proline methyl ester hydrochloride (**21**) using mixed anhydride activation gave the ester (**47**) in 83% yield, m.p. 116-8 °C. The α-bromo centre of (**47**) was found to be extremely reactive and the coupling reaction was carried out at a lower temperature than usual (-40 °C) and left stirring for only 30 min. Higher temperatures and longer reaction times resulted in epimerization at this chiral centre, detected from the appearance of a second set of signals in the ^{1}H and ^{13}C NMR spectra.

The reaction of α-bromo ester (**47**) with excess hydrazine hydrate gave the fused bicyclic compound (**49**) in 98% yield as an oil. The corresponding diastereomeric compound (**50**) was synthesized from (2*R*)-bromopropionic acid (**45**), using the procedure described above (Scheme 2. 9). The (*R,S*) α-bromo ester (**48**) was a crystalline compound and and its stereochemistry was unambiguously assigned from the results of a single crystal X-ray structure analysis (Fig. 2.10).

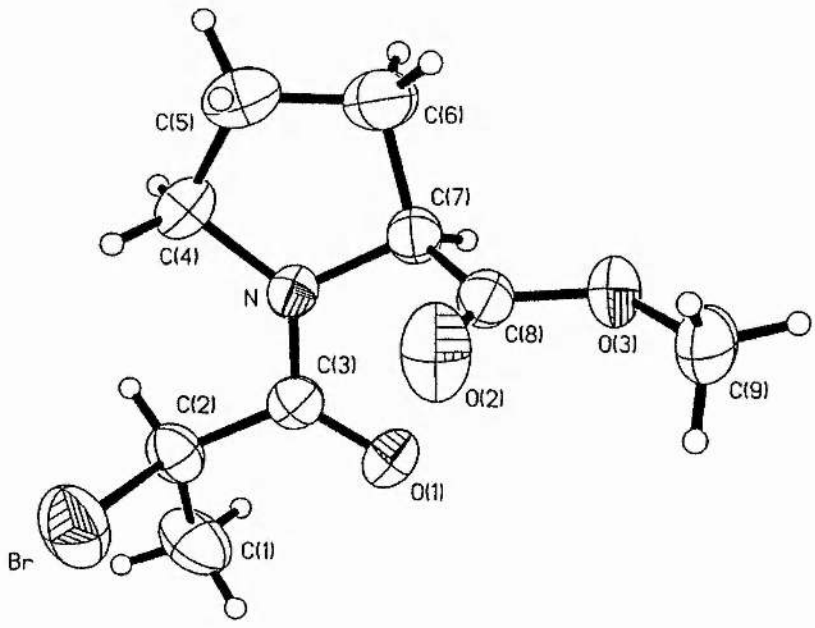
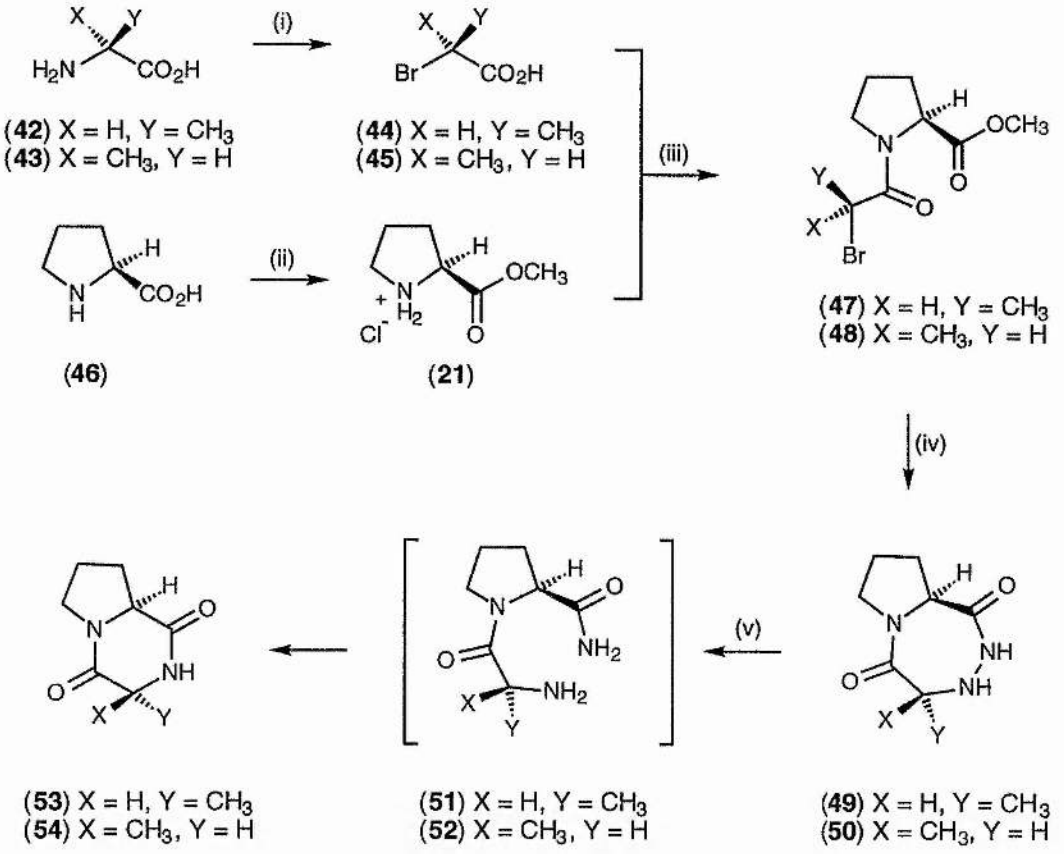


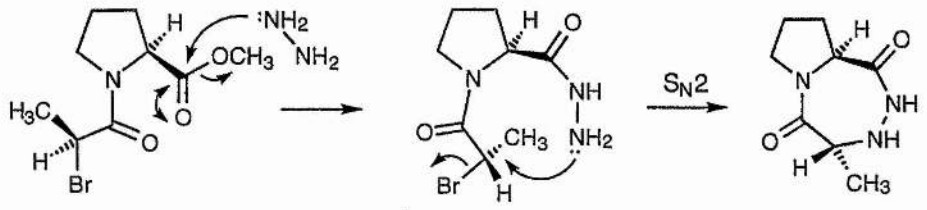
Figure 2.10: *X-ray structure of (48)*



Reagents and conditions: (i) HBr, KBr, $NaNO_2$, H_2O, 0 °C, 2 h; (ii) $SOCl_2$, MeOH, 65 °C, 30 min; (iii) NMM, *iso*-butyl chloroformate, THF/ DMF, -40 °C, 30 min; (iv) $NH_2NH_2.H_2O$, EtOH, 78 °C, 16 hr; (v) $NH_3(l)$, Na, -78 °C.

Scheme 2.9: *Synthesis and reactivity of bicyclic* cis *Ala-Pro mimetics (**49**) and (**50**)*

NMR studies showed that the cyclizations of (**47**) and (**48**) took 16 h to reach completion. The large difference in rate between the reaction of hydrazine hydrate with compounds (**47**) and (**48**), compared with the unsubstituted compound (**38**) obviously arises from the presence of a methyl group on the halogenated sp^3 carbon. This gives a strong indication of the mechanism of formation of the bicyclic compounds (Scheme 2.10). It can be assumed that the initial attack of the hydrazine occurs at the ester, giving an intermediate hydrazide. The rate determining step is probably the intramolecular S_N2 attack of the NH_2 group on the halogenated carbon to produce the cyclized product with inversion of configuration at the centre of attack.



Scheme 2.10: *Proposed mechanism of formation of the bicyclic hydrazides*

2.3.4 Reduction of the bicyclic hydrazides

The reductive cleavage of the N-N bond of the bicyclic hydrazides (**49**) and (**50**) proved more difficult than expected. The most usual literature method for achieving N-N cleavage is catalytic hydrogenolysis. However, the use of several catalysts, including Adams' catalyst, $PdCl_2$, palladium on charcoal and rhodium on alumina all failed to give the desired product. Alexakis[208] reported an ultrasound-assisted cleavage of hydrazine N-N bonds by Raney nickel[209] but this method also proved unsuccessful.

The dissolving metal reduction system has been used for the cleavage of N-N bonds where one or both nitrogens are attached to a carbonyl group.[210-12] This method was employed, using the procedure of Kemp *et al.*[210] to successfully reduce the bicyclic compound (**50**), giving one product in reasonable yield (Scheme 2.9).

It was clear, however, from spectroscopic data and from the chemical reactivity of the reduced compound that it was not the expected product (**52**):-

- The (EI) mass spectrum of the reduced product showed a molecular ion peak at *m/z* 168. This corresponded to loss of NH_3 from the expected product.

- The NMR spectra of the open chain compound (**52**) would be expected to show a mixture of *cis* and *trans* isomers. The ^{13}C NMR spectrum of the reduced compound showed the presence of one set of peaks only.

- The reduced product showed no reactivity when heated with acetic anhydride. The primary amino group of the open chain compound (**52**) would be expected to undergo acetylation under these conditions.

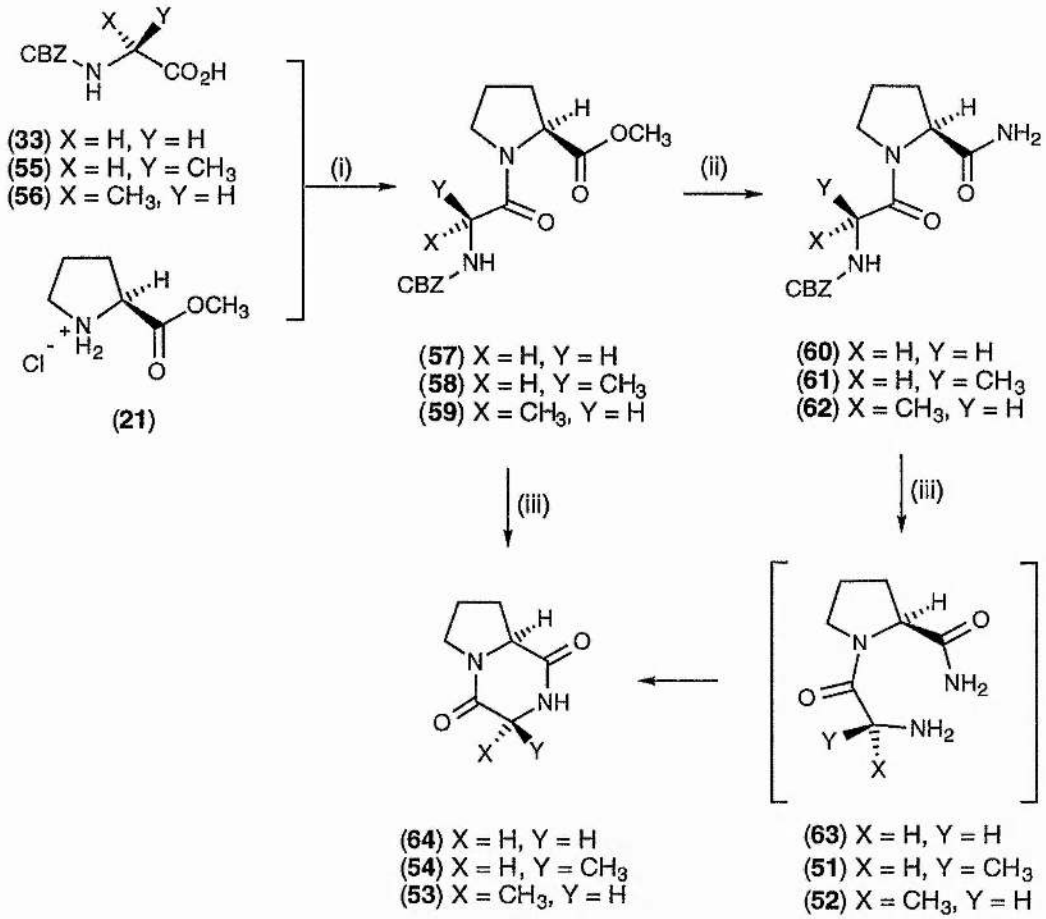
2.3.5 Alternative synthesis of the reduction products

In order to identify the products from the reduction of (**49**) and (**50**), we decided to prepare the *trans* isomers of the desired products (**51**) and (**52**) *via* a different route.

Both enantiomers of *N*-benzyloxycarbonyl alanine, (**55**) and (**56**), were coupled with proline methyl ester hydrochloride (**21**) using standard mixed anhydride methodology (Scheme 2.11). The amino-protected esters, (**58**) and (**59**) were obtained in 84% and 72% yield respectively. Removal of the CBZ protecting group by catalytic hydrogenation gave the known cyclic compounds, (**53**) and (**54**) in excellent yield {for compound (**53**), m.p. 128-30 °C (lit.,[203] 127-9 °C)}. The intramolecular reaction of an amine with an ester to form a six-membered ring is facile, and is the standard method for the preparation of diketopiperazines.

Ammonolysis of the dipeptide esters, (**58**) and (**59**) gave the dipeptide amides, (**61**) and (**62**) as crystalline solids (for compound (**61**), m.p. 167-9 °C). Removal of the CBZ protecting group again led to the formation of the individual diastereomeric diketopiperazines, presumably *via* an intramolecular transamidation reaction (see section 2.4). It was surprising that the formation of the diketopiperazines was equally facile to the analogous reaction with the methyl esters (**58**) and (**59**) when the leaving group was that of an amide.

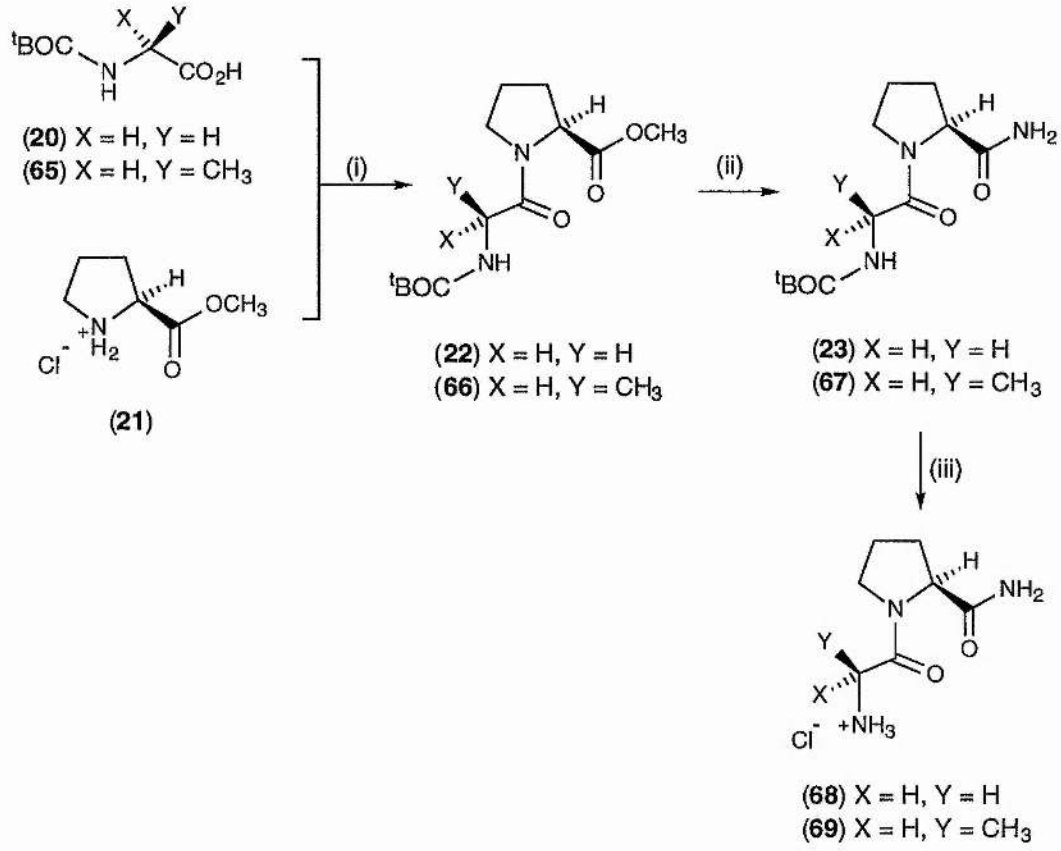
The CBZ groups were removed in an acidic solution, in an attempt to trap out (**51**) and (**52**) as their hydrochloride salts before cyclization occurred. Although this was successful and hindered the cyclization process, ^{1}H NMR spectra showed that after 8 h, >80% of the deprotected dipeptide had formed the diketopiperazine.



Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 2 h; (ii) $NH_3(g)$, MeOH, rt., 48 h; (iii) $H_2(g)$, Pd/C(cat), MeOH, rt., 8 h.

Scheme 2.11: *Alternative diketopiperazine syntheses*

An alternative method of trapping out the dipeptide amides directly would be to choose an acid labile protecting group for the amine, such as ^{t}BOC. The ^{t}BOC protected dipeptides, (**23**) and (**67**) were therefore prepared as outlined in Scheme 2.12. Removal of the ^{t}BOC groups using hydrogen chloride gas gave the dipeptide amide hydrochlorides (**68**) and (**69**) as hygroscopic solids. The ^{1}H and ^{13}C NMR spectra for each compound contained two sets of peaks, the minor set due to the *cis* isomer (15%), providing strong evidence in support of open chain structures.



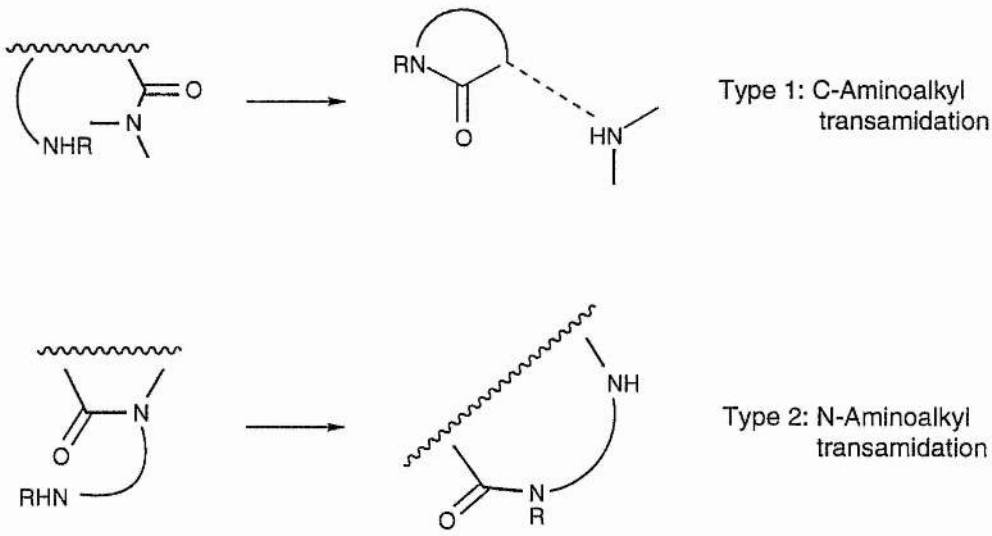
Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 2 h; (ii) NH_3(g), MeOH, rt., 48 h; (iii) HCl(g), EtOAc, 0 °C, 15 min.

Scheme 2.12: *Synthesis of open chain dipeptide amides*

2.4 Intramolecular transamidation reactions

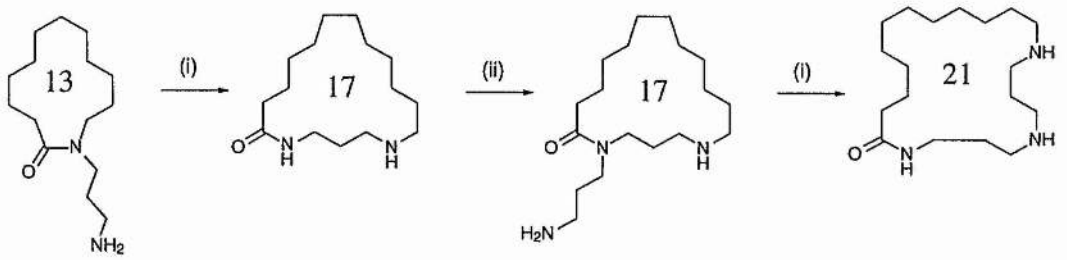
2.4.1 Introduction

There are very few cases reported in the literature of intramolecular transamidation reactions, and even fewer examples where such reactions occur spontaneously.[213] The few existing examples have been classified into two types (Scheme 2.13).[214,215] Type 1 is an amine elimination which occurs with formation of a cyclic lactam and the second type is a rearrangement, combined with N→N′ acyl migration.



Scheme 2.13: *Types of intramolecular transamidation*

An example of Type 2 is the "zip" reaction, an interesting ring expansion which proceeds *via* intramolecular transamidation (Scheme 2.14).[216,217] Starting from *N*-substituted lactams, successive incorporation of aminopropyl units produces macrocyclic polyazalactams. The zip reaction can be used to form macrocycles with up to 53 ring atoms. These reactions are not spontaneous and strong bases such as ^{t}BuOK or potassium-3-aminopropyl amide (KAPA) are required to deprotonate the exo-cyclic amine group.

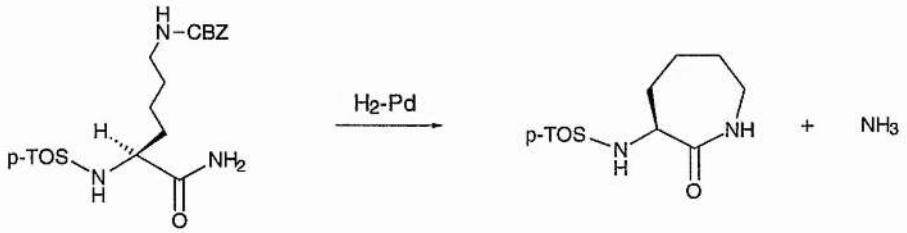


Reagents: (i) KAPA, H_3O^+; (ii) Na salt of the amide, CH_2=CH-CN; H_2/Pt.

Scheme 2.14: *The "zip" reaction*

The reaction is only successful with aminopropyl groups, as this chain length results in a favourable six-membered transition state. In such cases, there is no obvious release of ring strain and the driving force of the reaction can be regarded as the conversion of the initially generated amino anion into the comparatively stable amide anion.

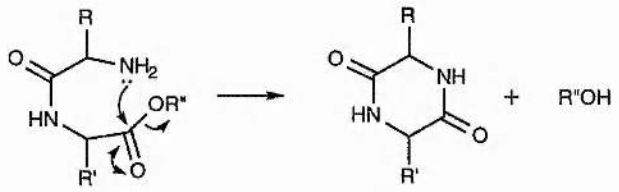
Examples of spontaneous intramolecular transamidation tend to be of type 1, where there is an additional entropic driving force due to the elimination of an amine molecule. It was noted by Barrass and Eltmore in 1957,[218] that the side chains of derivatives of basic amino acids such as lysine, underwent spontaneous reactions of type 1 to produce cyclic lactams (Scheme 2.15).



Scheme 2.15: *Formation of cyclic lactams from basic amino acid amides*

2.4.2 Formation of diketopiperazines

A problem often encountered in both solution and solid phase peptide synthesis is that of base catalyzed ring closure. Such side reactions often occur during the process of adding a third residue to a dipeptide derivative. In particular, alkyl esters of dipeptides are known to undergo spontaneous cyclization to produce 2,5-diketopiperazines (DKPs) (Scheme 2.16).



Scheme 2.16: *Formation of DKPs from dipeptide esters*

Both amide bonds in DKPs are in the *cis* configuration. Accordingly, the cyclization process requires the energetically unfavourable conversion of the *trans* open chain dipeptide to the less stable *cis* form. The energy liberated in the formation of the very stable DKP structure provides the necessary driving force for the reaction. There are certain cases in which spontaneous DKP formation is especially favoured.

- Ring formation is particularly pronounced in glycine-containing peptides, since the absence of a bulky side chain makes the cyclization process easier.[182]

- DKP formation is highly favoured for proline-containing peptides as the cyclic side chain lies in the plane of the DKP molecule and is out of the way during ring closure. In addition, the *cis* conformation (which is required for cyclization) is more stable relative to the *trans* than for other amino acids (see Section 1.2.3, p. 24).[81-84]

- DKPs are more readily formed when one of the residues has (*S*) chiral configuration and the other, (*R*). In such cases, the side chains are on opposite sides of the general plane of the DKP.[203] If both amino acids have the same configuration, then both side chains are on the same side and their bulk interferes with ring closure.

DKP formation is known to occur readily in methyl, ethyl and benzyl esters of dipeptides, but not in tertiary butyl esters. The formation of diketopiperazines from

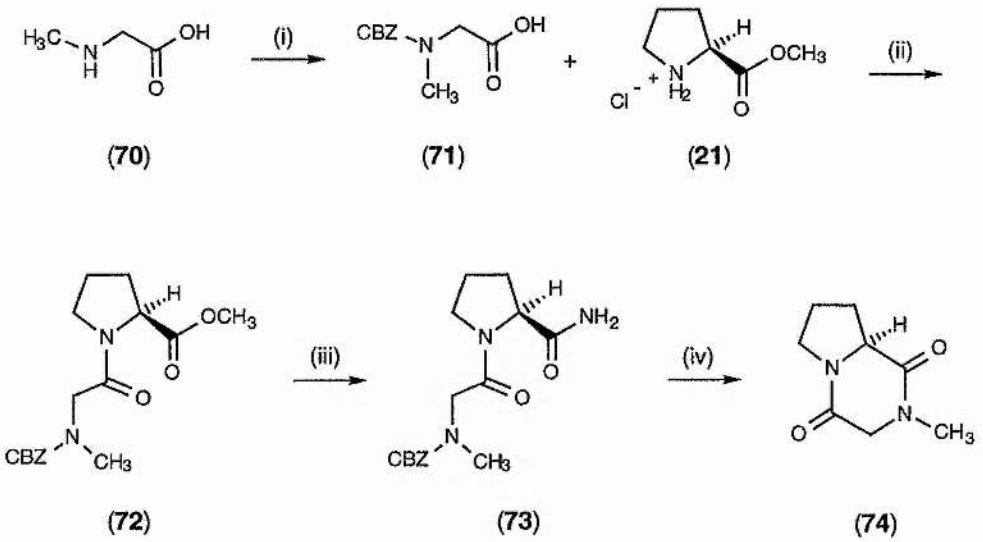
dipeptide amides was first reported by Huang and Niemann in 1950,[219] who observed this phenomenon with non-prolyl dipeptides and suggested the reaction was restricted to primary amides.

2.4.3 Diketopiperazine formation from X-Pro amides

The facile transformation of the dipeptide amides, (**51**), (**52**) and (**63**) into the DKPs, (**53**), (**54**) and (**64**) was an undesirable reaction for the purpose of obtaining *cis* X-Pro peptides (Scheme 2.11, p. 78). We therefore decided to investigate this intramolecular transamidation further to find out how it could be prevented.

Substitution of the amine nitrogen was first tried to see if this would slow the rate of cyclization (Scheme 2.17). Due to the high cost of *N*-methyl alanine, sarcosine (*N*-methyl glycine) (**70**), was employed. *N*-Benzyloxycarbonyl sarcosine (**71**) was coupled to proline methyl ester hydrochloride (**21**) using mixed anhydride methodology to produce the protected dipeptide ester (**72**) (Scheme 2.17). Compound (**72**) was isolated as an oil and was purified using column chromatography. Ammonolysis of (**72**) in methanol produced the dipeptide amide (**73**) in almost quantitative yield. The NMR spectra of (**72**) and (**73**) were extremely complex, showing the presence of four conformational isomers. This is due to the sarcosyl tertiary amide bond which exists as both *cis* and *trans* conformations (see section 1.3.4) in a roughly 4:6 (*cis*:*trans*) ratio. Since both conformations of the prolyl amide bond were also present in an approximately 1:4 (*cis*:*trans*) ratio, the result was a mixture of four conformers in an 8:12:32:48 ratio.

After removal of the CBZ group by hydrogenolysis, both the ^{1}H and ^{13}C NMR spectra of the product were greatly simplified, showing only one set of peaks. This suggested that the *N*-methyl DKP (**74**) had been formed. Although this compound has been reported in the literature as an oil, no accompanying spectroscopic data were published.[220] Further spectroscopic and analytical analysis confirmed the structure of the DKP to be that of (**74**). Thus, it was clear that substitution of the amine nitrogen did not significantly slow down the transamidative cyclization reaction.



Reagents and conditions: (i) NaOH (1 mol dm^{-3}), $PhCH_2OCOCl$, 0 °C, 4 h, 96%; (ii) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 2 h, 77%; (iii) $NH_3(g)$, MeOH, rt., 48 h, 80%; (iv) $H_2(g)$, Pd/C, MeOH, rt., 12 h, 90%.

Scheme 2.17: *Synthesis of an* N*-methylated DKP compound*

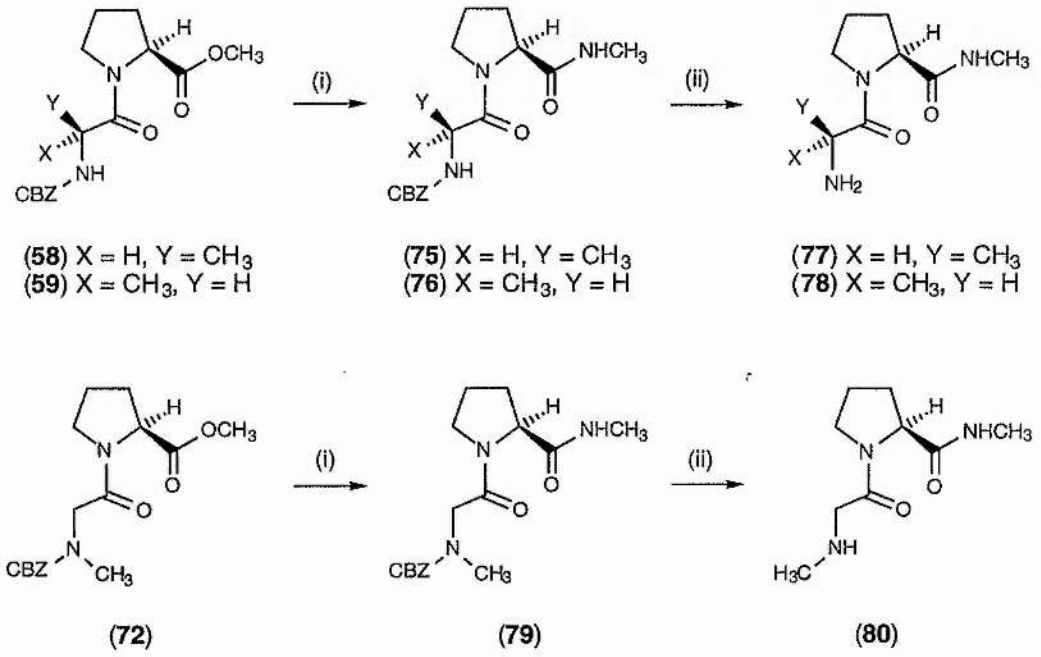
Secondly, the amide functionality was changed to see if this would have any effect on the rate of cyclization. The dipeptide esters, (**58**), (**59**) and (**72**), derived from (*S*)-alanine, (*R*)-alanine and sarcosine respectively, were dissolved in methanol saturated with methylamine gas and the reaction was followed by TLC and by NMR spectroscopy (Scheme 2.18). The rate of the reaction of methylamine with alkyl esters was found to be faster than that of ammonia and the ^1H NMR spectra showed the disappearance of the methyl ester (δ 3.4 ppm) after 6 h. This greater reactivity was expected, due to the increased nucleophilicity of methylamine and also due to its greater solubility and therefore higher concentration, in methanol.

The resulting dipeptide methylamides, (**75**), (**76**) and (**79**) were then hydrogenated and the ^1H NMR spectra showed the complete disappearance of aromatic signals (δ 5.0 & 7.5 ppm) after 12 h. The structures of the resulting compounds, (**77**), (**78**) and (**80**) were assigned on the basis of spectroscopic evidence.

- Two sets of signals were observed in the ^{13}C NMR spectra of the products. This characteristic was suggestive of *cis/trans* isomerism and therefore indicative of open chain structures.

- The products reacted easily with acetic anhydride and NMR studies showed the appearance of new signals at 2.1 ppm, due to *N*-acetyl functionalities. These observations provide good evidence for the presence of a free amine group.

- After three days, NMR spectra recorded for the (*R,S*) diastereomer (**78**) showed a new set of minor peaks. These corresponded exactly to the signals observed for the diketopiperazine, (**53**). No set of minor peaks were observed in NMR spectra of the (*S,S*) diastereomer, (**77**) after the same time period had elapsed. These observation are consistent with the greater propensity of (*R,S*) dipeptides to undergo DKP formation (see p. 82).



(**58**) X = H, Y = CH_3
(**59**) X = CH_3, Y = H

(**75**) X = H, Y = CH_3
(**76**) X = CH_3, Y = H

(**77**) X = H, Y = CH_3
(**78**) X = CH_3, Y = H

(**72**) (**79**) (**80**)

Reagents and conditions: (i) $H_2NCH_3(g)$, MeOH, rt., 8 h; (ii) $H_2(g)$, Pd/C, MeOH, rt., 12 h.

Scheme 2.18: *Effect of amide substitution on transamidation*

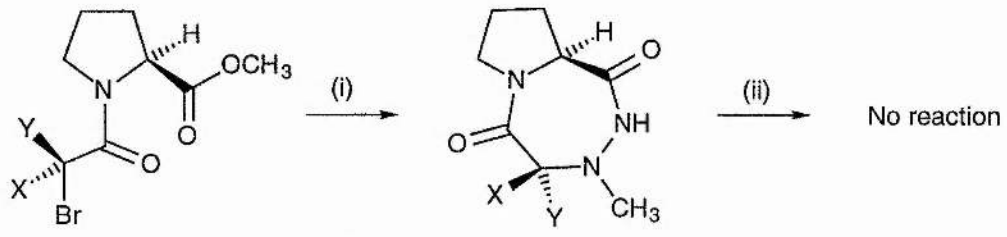
From these observations, it was clear that substitution of the primary amide for a methylamide group led to a dramatic reduction in the rate of DKP formation. In order to synthesize *cis* prolyl peptides *via* reductive N-N bond cleavage, a secondary amide had to be produced in order to prevent consequent deamination of the product. These results also showed that, in contradiction to the observations of Huang and Niemann,[219] DKP formation can occur for secondary amides, albeit at a much slower rate.

2.5 Synthesis of *N*-substituted *cis* prolyl peptides

2.5.1 Introduction

The next step towards the synthesis of *cis* X-Pro peptides was therefore to use substituted hydrazines in the formation of 5,7-bicyclic compounds. It was hoped that the N-N bond in these compounds would be cleaved under the same conditions as (**49**) and (**50**) but that the resulting open chain peptides would not undergo such rapid diketopiperazine formation.

In order to establish the regioselectivity of the ring formation with substituted hydrazines, the α-bromo-dipeptide esters, (**38**), (**47**) and (**48**) were reacted with methylhydrazine (**84**), the simplest monosubstituted hydrazine (Scheme 2.19). Excess methylhydrazine was added to the α-bromo dipeptide esters in refluxing ethanol and TLC indicated the complete disappearance of the starting material after around 15 min.



(**38**) X=H, Y=H
(**47**) X=H, Y=CH_3
(**48**) X=CH_3, Y=H

(**81**) X=H, Y=H
(**82**) X=H, Y=CH_3
(**83**) X=CH_3, Y=H

Reagents and conditions: (i) $NHCH_3NH_2$, EtOH, 78 °C, 30 min; (ii) Na, NH_3(l), THF, -60 °C, 1 h.

Scheme 2.19: *Synthesis and reactivity of* N-*methylated bicyclic* cis *X-Pro mimetics*

The NMR spectra of the products were indicative of bicyclic compounds since no *cis/trans* conformational isomerism was observed. The reactions were apparently highly regio and stereoselective in all cases and NMR spectroscopy indicated that one product had formed in greater than 95% yield. Although all analytical and spectral

data were consistent with the structures, (**81**), (**82**) and (**83**), the orientation of the methylhydrazine moiety in the products was not immediately clear. Since it had been understood that the ester functionality in (**38**), (**47**) and (**48**) was primarily attacked by unsubstituted hydrazine (Scheme 2.10), it was assumed that attack by methylhydrazine would initially occur at the same site. It is known, however, that methylhydrazine generally reacts with simple esters to give a mixture of 1-acyl-1-methyl and 1-acyl-2-methyl isomers.[221]

2.5.2 Structural studies

Various NMR experiments were employed to try and establish whether the methyl group was situated on the amide or the amine nitrogen However, these proved inconclusive and a single crystal X-ray diffraction study of compound (**81**) was carried out in order to determine the correct structure.

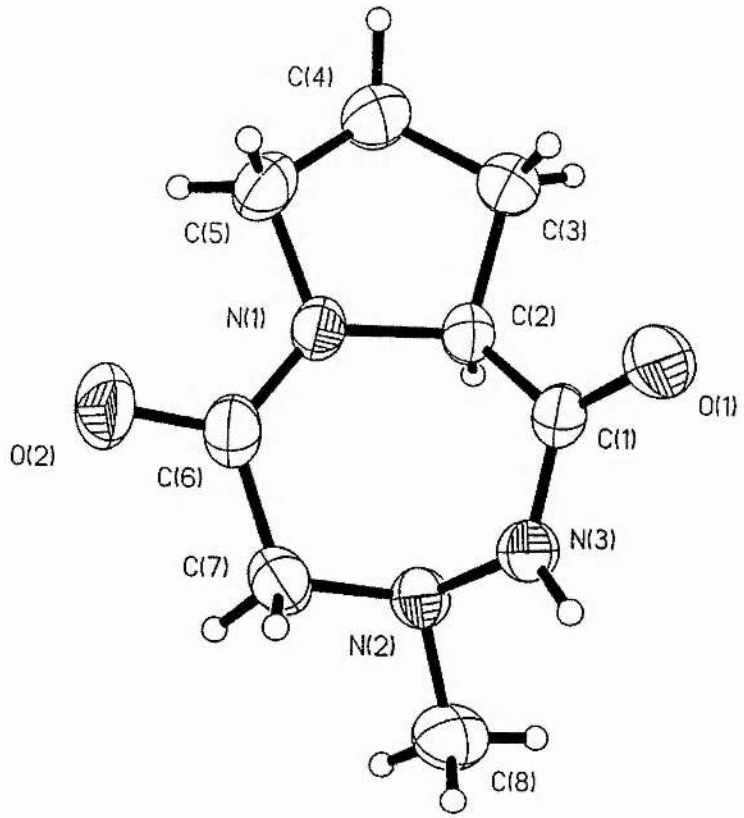


Figure 2.11: *X-ray structure of (81)*

The X-ray structure (Fig. 2.11) showed that the methyl group was situated on the amine nitrogen. In the solid state, the seven-membered triazepine ring was found to be in a pseudo boat conformation with the *N*-methyl group in an equatorial position (Fig. 2.12). The pyrrolidine ring has an envelope structure with approximate C_s symmetry, the mirror plane passing through the C(4) or C$^{\gamma}$ atom. The N(1), C(2), C(3) and C(5) groups are almost planar and the C(4) and C(1) atoms are on opposite sides of the plane. Thus, the conformation of the ring is C_s-C$^{\gamma}$-*exo* as defined by Ashido and Kakudo[222] and corresponds to conformation A according to Balasubramian *et al.*[223]

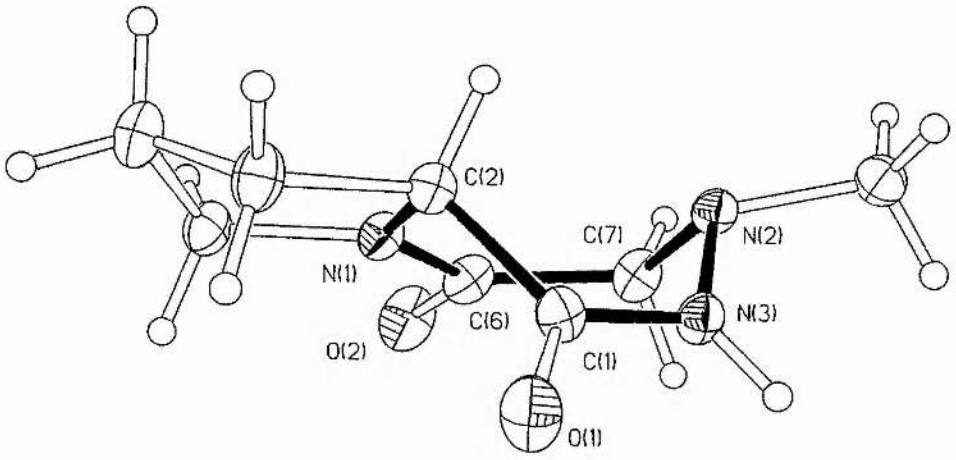


Figure 2.12: *X-ray structure of (81)*

The rapid formation of compounds (**81**), (**82**) and (**83**) supports the previously proposed mechanism (Scheme 2.10). The increased nucleophilicity of the methylated nitrogen would facilitate the intramolecular cyclization step. If this step proceeded by an S_N2 mechanism, then inversion of stereochemistry would have taken place at the α carbon.

In order to establish the correct stereochemistry of the products, a single crystal X-ray diffraction study was carried out on the product of the reaction between the (*S,S*) α-bromo ester (**47**) and methylhydrazine. The product bicyclic compound (**82**) was found to posess (*R,S*) stereochemistry (Fig. 2.13). This showed that inversion of configuration had taken place and provides evidence for an S_N2-type mechanism.

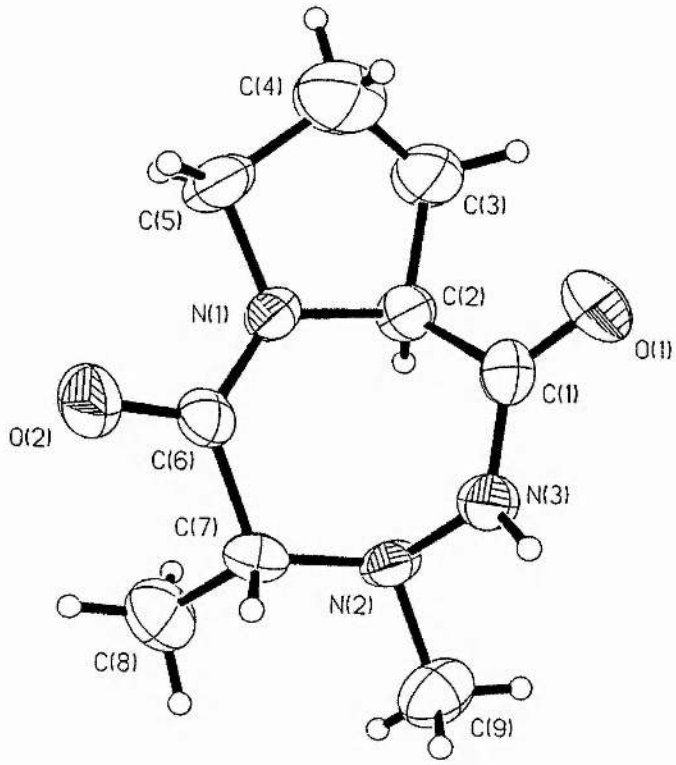


Figure 2.13: *X-ray structure of (82)*

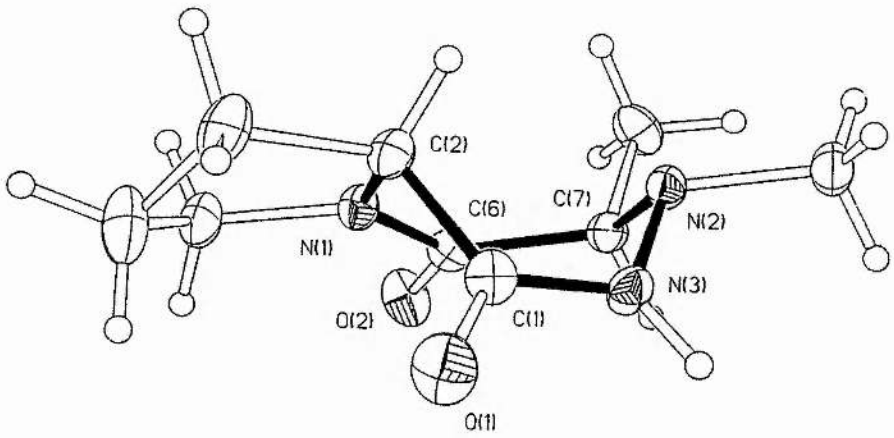


Figure 2.14: *X-ray structure of (82)*

The seven-membered ring of (**82**) adopts a pseudo boat configuration (Fig. 2.14), as observed for the related compound (**81**). In this case, however, the conformation of the pyrrolidine ring is C_s-Cγ-*endo* as defined by Ashido and Kakudo[222] and corresponds to conformation B according to Balasubramian *et al.*[223]

The *N*-methylated bicyclic compounds, (**81**), (**82**) and (**83**), were all found to be stable to dissolving metal reduction (Scheme 2.19). Addition of excess amounts of sodium and long reaction times failed to cause reaction of the starting materials. The reasons for this are not immediately clear. The presence of the alkyl group on the amine nitrogen obviously deactivates the compounds relative to the unsubstituted analogues, (**41**), (**50**) and (**51**). This could be due to steric constraints, as the methyl group may make it harder for a solvated electron to attack the amide carbonyl.

2.5.3 Synthetic strategy

From the results obtained from the preceding studies, it was clear that *cis* X-Pro peptides could only be prepared from compounds which possessed the following characteristics. Firstly, a bicyclic precursor was required which did not contain an alkyl group on the amine nitrogen of the hydrazide bond. This appeared to prevent successful reductive cleavage of the N-N bond using the sodium/ liquid ammonia methodology (Scheme 2.19). Secondly, the presence of an alkyl group on the amide nitrogen of the hydrazide bond was essential for the prevention of diketopiperazine formation (see section 2.4.3).

Previous studies had shown that the unsubstituted nitrogen of monoalkylated hydrazines reacted in preference with the ester groups of α-bromo dipeptide esters such as (**37**), (**47**) and (**48**) (Scheme 2.19). It was therefore clear that a hydrazine derivative was required where the unsubstituted nitrogen was deactivated relative to the alkylated nitrogen. Protection of the unsubstituted nitrogen was not straightforward due to the unusual reactivity of hydrazines.

2.5.4 Reactivity of hydrazines

In contrast to the situation in amines, substitution of hydrogens by alkyl groups in hydrazines lowers the basicity. The hydrazine, however, is protonated on the more highly alkylated nitrogen. Similarly, alkylation of hydrazines usually takes place on

the more substituted nitrogen.[224] Acylation of unsymmetrical hydrazines, however, is less predictable than alkylation and more dependent on the conditions used.[221]

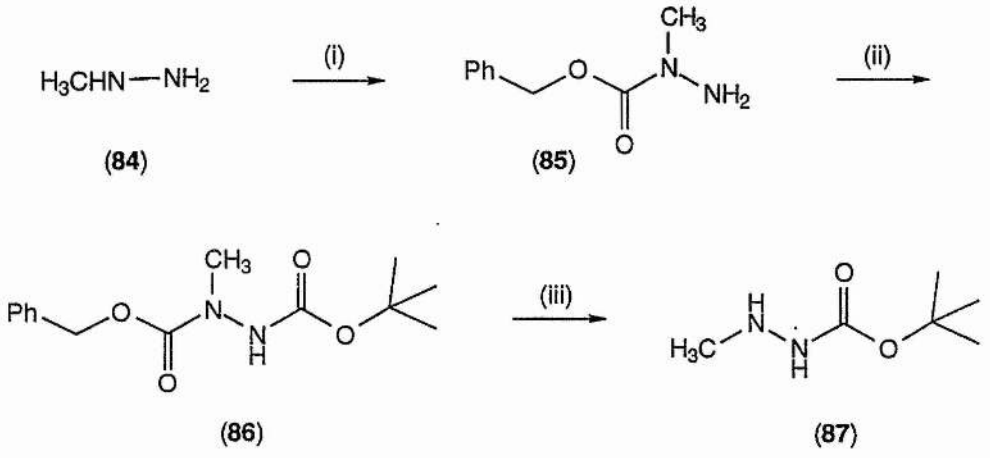
The methyl-bearing nitrogen of methylhydrazine (**84**) has been shown to attack selectively reactive functional groups such as carboxylic acid anhydrides and chlorides.[225,226] Conditions have been devised by Peet *et al.* for acetylating methylhydrazine using acetyl chloride to give 1-methyl-1-acetylhydrazine in excellent yield.[227] In contrast, the reaction of an ester and methylhydrazine yields a mixture of the 1-acyl-2-methyl and 1-acyl-1-methyl isomers.[221] It is thus extremely difficult to acylate selectively the primary amine of methylhydrazine in one step.

We decided that the most successful strategy towards the synthesis of a *cis* Gly-Pro mimetic would therefore require reaction of the α-bromo ester (**38**) with a suitably N-protected-N′-methylhydrazine, followed by subsequent deprotection and cyclization (Scheme 2.21).

An N-protected-N′-methylhydrazine could easily be obtained, starting with selective protection of the methyl-bearing nitrogen. Protection of the primary amine then with a different protecting group which was stable to conditions for the deprotection of the methylated nitrogen would give the desired compound. We decided to use the benzyloxycarbonyl group to protect the latter, since this group is easily removed by catalytic hydrogenation which would not affect a *tert*-butoxycarbonyl group on the adjacent nitrogen.

Pederson reported the synthesis of *N*-benzyloxycarbonyl-*N*-methylhydrazine (**85**) in 60% yield from the reaction of methylhydrazine with dibenzylcarbonate.[228] However, as the procedure involved separating the required compound from the 1-acyl-2-methyl isomer, we opted for an alternative synthesis. By using benzyl chloroformate as the acylating agent, and employing the conditions of Peet *et al.*,[227] (**85**) was obtained in high yield (89%) with excellent regioselectivity. The ^{1}H NMR showed that less than 3% of the undesired isomer had been formed.

The monoacylated derivative of methylhydrazine was then reacted with ditertiarybutylcarbonate to give the dihydrazide (**86**). Removal of the CBZ group by hydrogenolysis using palladium on activated charcoal gave the desired product, *N-tert*-butoxycarbonyl-*N′*-methylhydrazine (**87**), m.p. 46-9 °C.

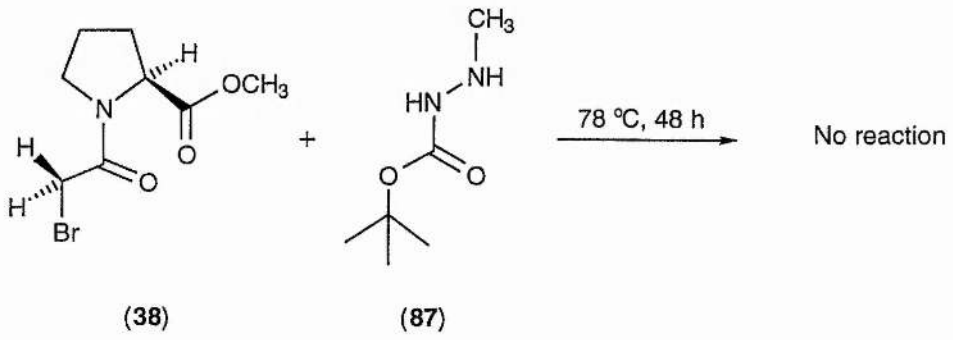


Reagents and conditions: (i) $PhCH_2OCOCl$, CH_2Cl_2, 0 °C, 30 min, 89%; (ii) $(^tBuO_2C)_2O$, ^iPrOH/ CH_2Cl_2, rt., 3 h, 80%; (iii) $H_2(g)$, Pd/C, MeOH, rt., 12 h, 85%.

Scheme 2.20: *Synthesis of* N-tert-*butoxycarbonyl*-N′-*methylhydrazine*

2.5.5 Synthesis of *N*-methyl *cis* X-Pro mimetics

The methylhydrazine derivative (**87**) and the dipeptide ester (**38**) failed to react together in boiling ethanol, even after 48 h. This was probably due to steric hindrance since the synthesis of secondary and tertiary amides from carboxylic esters is known to be difficult for this reason.[229]

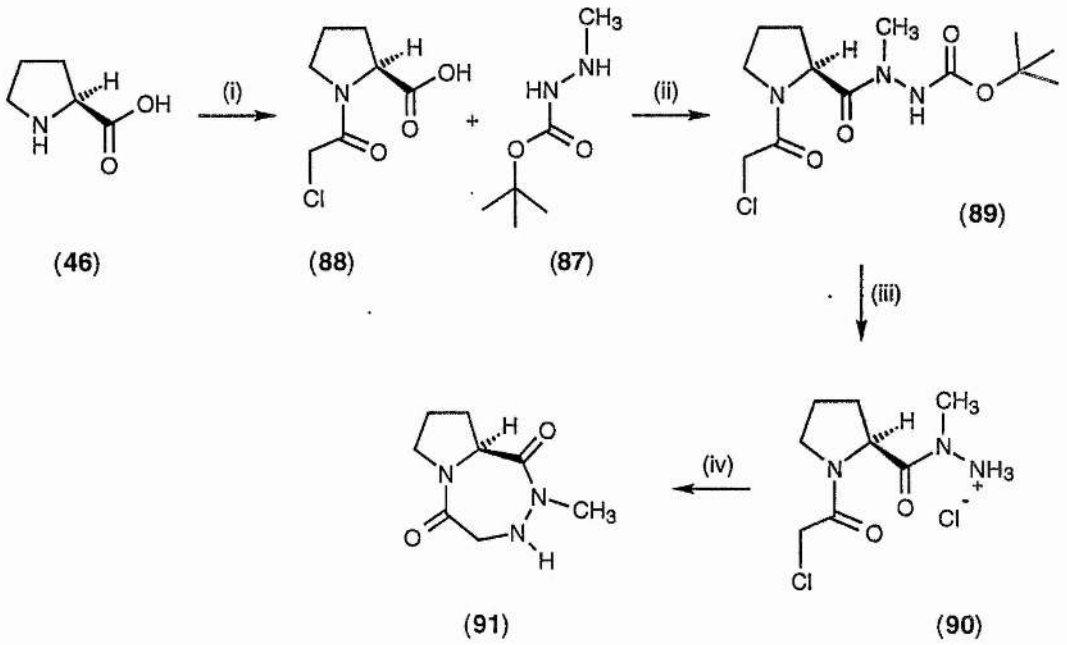


Scheme 2.21: *Attempted reaction of (87) with dipeptide ester (38)*

Although the reaction of (**87**) with esters was not favourable, such hindered amines are known to react with carboxylic acid groups using various peptide coupling techniques.[193] *N*-Chloroacetyl-(2*S*)-proline (**88**) was prepared from chloroacetyl chloride and proline (**46**) by the method of Ronwin[230] (Scheme 2.22) and reacted with the methylhydrazine derivative (**87**). When the mixed anhydride method was used, compound (**89**) was obtained in 79% yield, m.p. 64-7 °C. Although the coupling reagent PyBroP,[192] which is supposed to be a specialized reagent for the coupling of *N*-methyl amino acids, gave a slightly improved yield (86%), its considerable expense did not justify its use.

Removal of the ^{t}BOC protecting group using HCl gave the hydrochloride salt (**90**) as a hygroscopic white solid. ^{1}H NMR showed the complete loss of the *tert*-butyl signal at 1.48 ppm. Excess NMM was then added to a solution of (**90**) in methanol and after removal of the solvent, the reaction mixture was partitioned between water and ethyl acetate. The organic layer was found to contain the bicyclic product (**91**), m.p. 75-9 °C.

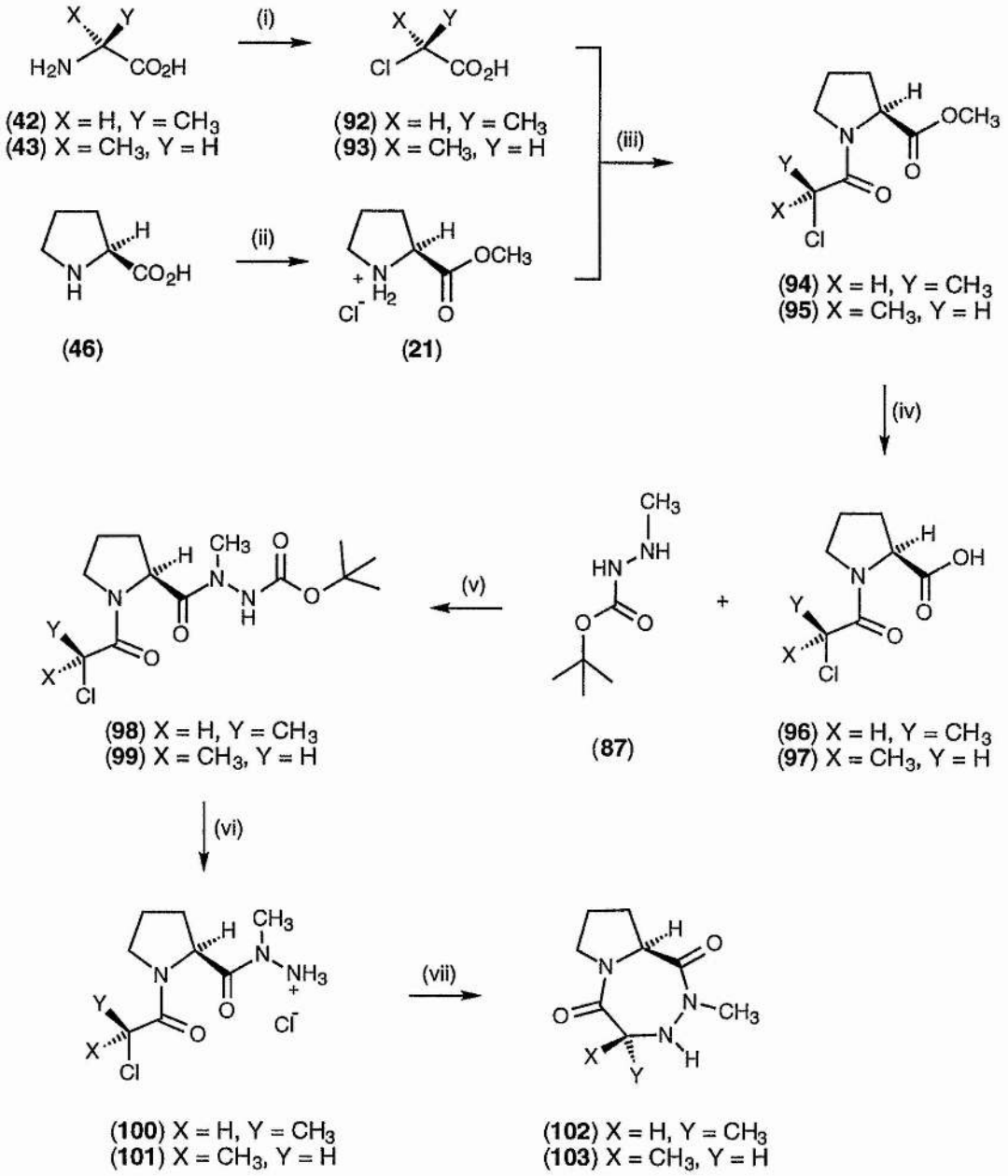
The intramolecular S_N2 cyclization step appeared to be extremely facile, even though the leaving group in this case was chloride rather than bromide, as in the formation of (**81**) (Scheme 2.19). Although the mass spectrum of (**91**) gave a molecular ion at the same *m/z* as compound (**81**), all other spectroscopic data were different, confirming it as a chemically distinct isomer. Unlike that of the isomeric compound (**81**), the ^{13}C NMR spectrum of (**91**) contained two sets of peaks. This is presumably due to conformational isomerism, as displayed by the related *cis* Gly-Pro bicyclic compound, (**39**).



Reagents and conditions: (i) $ClCH_2COCl$, $NaHCO_3$, CH_2Cl_2, 0 °C, 30 min, 68%; (ii) NMM, *iso*-butyl chloroformate, THF, -15 °C, 12 h, 73%; (iii) HCl(g), EtOAc, 0 °C, 15 min, 100%; (iv) NMM, MeOH, 5 min, 79%.

Scheme 2.22: *Synthesis of an* N*-methylated* cis *Gly-Pro mimetic*

The synthesis of analogous *cis* (*R*) and (*S*) Ala-Pro mimetics was more difficult due to unwanted reactions at the second chiral centre. The carboxylic acid precursors, (**96**) and (**97**) were obtained from hydrolysis of esters (**94**) and (**95**) respectively, using 1 mol dm^{-3} sodium hydroxide (Scheme 2.23). It was this hydrolysis step which rendered use of the bromo dipeptide esters, (**47**) and (**48**) unsuitable. The brominated carbon centre in these compounds was found to be extremely reactive (see section 2.4.3). Although the ester groups of (**47**) and (**48**) were readily hydrolysed, ^{1}H and ^{13}C NMR spectra showed the presence of two sets of signals. We therefore reasoned that the brominated centre had undergone base-catalyzed enolization which led to scrambling of the stereochemistry.



Reagents and conditions: (i) HCl, KCl, $NaNO_2$, H_2O, 0 °C, 2 h, 91%; (ii) $SOCl_2$, MeOH, 65 °C, 30 min; (iii) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 1.5 h; (iv) 1 mol dm^{-3} NaOH, MeOH/ H_2O, rt., 2 h; (v) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 12 h; (vi) HCl(g), EtOAc, 0 °C, 15 min; (vii) NMM, MeOH, 5 min.

Scheme 2.23: *Synthesis of* N*-methyl* cis *Ala-Pro mimetics*

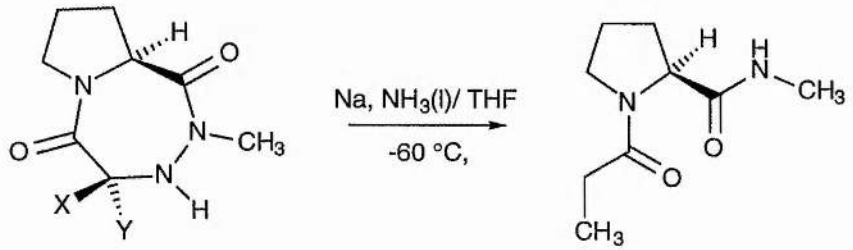
The α-chloro analogues, (**94**) and (**95**) were found to be less reactive and underwent base-catalyzed hydrolysis without any epimerization. The (2*S*) and (2*R*) chloropropionic acids, (**92**) and (**93**) were synthesized according to the method of Greenstein *et al.*[231] and coupled with proline methyl ester (**21**) using mixed anhydride methodology. The esters (**94**) and (**95**) were obtained in 71% and 83% yield respectively (for compound (**95**), m.p. 118-20 °C). Base catalyzed hydrolysis of (**94**) and (**95**) then gave the corresponding acids, (**96**) and (**97**) (for compound (**96**), m.p. 164-5 °C). The ^{1}H and ^{13}C NMR spectra showed only one set of peaks for each compound, suggesting that no side reactions had taken place.

The acids, (**96**) and (**97**) were then coupled to the methylhydrazine derivative (**87**) using mixed anhydride methodology to produce the prolyl hydrazides, (**98**) and (**99**) in reasonable yield (62% and 66% respectively), (Scheme 2.23). These compounds were easily purified by column chromatography and showed the expected spectroscopic and analytical data (for compound (**98**), m.p. 81-3 °C and for (**99**), m.p. 132-5 °C). Removal of the ^{t}BOC groups, followed by the subsequent cyclization of the resulting hygroscopic salts, (**100**) and (**101**), using 2 equivalents of NMM gave (**102**) and (**103**) (for compound (**102**), m.p. 108-10 °C and for (**103**), m.p. 112-15 °C). The bicyclic compounds, (**102**) and (**103**) were found to have different spectroscopic characteristics to the isomeric compounds (**82**) and (**83**) (Scheme 2.19).

2.5.6 Reductive cleavage of bicyclic compounds

Compounds (**82**) and (**83**) were found to be completely stable to reduction by sodium in liquid ammonia (Scheme 2.19). In contrast, compounds (**102**) and (**103**) were easily reduced under these conditions. Although the sodium in liquid ammonia reducing system is a general method for the cleavage of N-N bonds,[210-12] the mechanism of cleavage has never been investigated.

An unexpected product was obtained from the reaction of (**102**) and (**103**) with sodium in liquid ammonia at -60 °C (Scheme 2.24). The first surprising observation was that both diastereomers gave the same reduction product. This indicated loss of one of the chiral centres. Secondly, the ^{1}H NMR spectra contained a triplet at 1.1 ppm and a quartet at 2.3 ppm which is characteristic of the presence of an ethyl group.



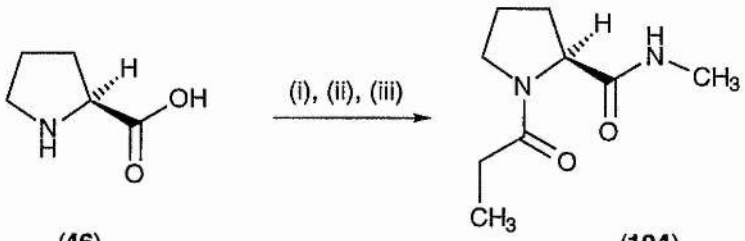
(**102**) X = H, Y = CH_3
(**103**) X = CH_3, Y = H

(**104**)

Scheme 2.24: *Reduction of bicyclic hydrazides (102) and (103)*

It seemed that as well as N-N bond cleavage and loss of ammonia, a C-N bond had been cleaved to give the reduction product (**104**), m.p. 49-52 °C. Compound (**104**) was prepared by an alternative route (Scheme 2.25) and all spectral and analytical data matched those of the reduction product of (**102**) and (**103**).

In order to make sure that (**104**) was not the result of over-reduction of the desired dipeptide products, the reaction was repeated with (**102**), using 0.5 equivalents of sodium. NMR spectra clearly showed the presence of (**104**) as well as some unreacted starting material. This demonstrated that a rapid C-N hydrogenolytic fission had taken place in preference to the N-N reductive cleavage. This event was very rapid and could not be prevented by altering the reaction conditions.

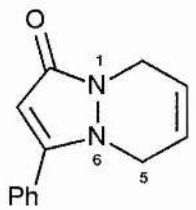


(46) (104)

Reagents and conditions: (i) CH_3CH_2COCl, $NaHCO_3$, CH_2Cl_2, 0 °C, 30 min, 82%; (ii) $SOCl_2$, MeOH, 65 °C, 30 min, 100%; (ii) H_2NCH_3(g), MeOH, rt., 8 h, 99%.

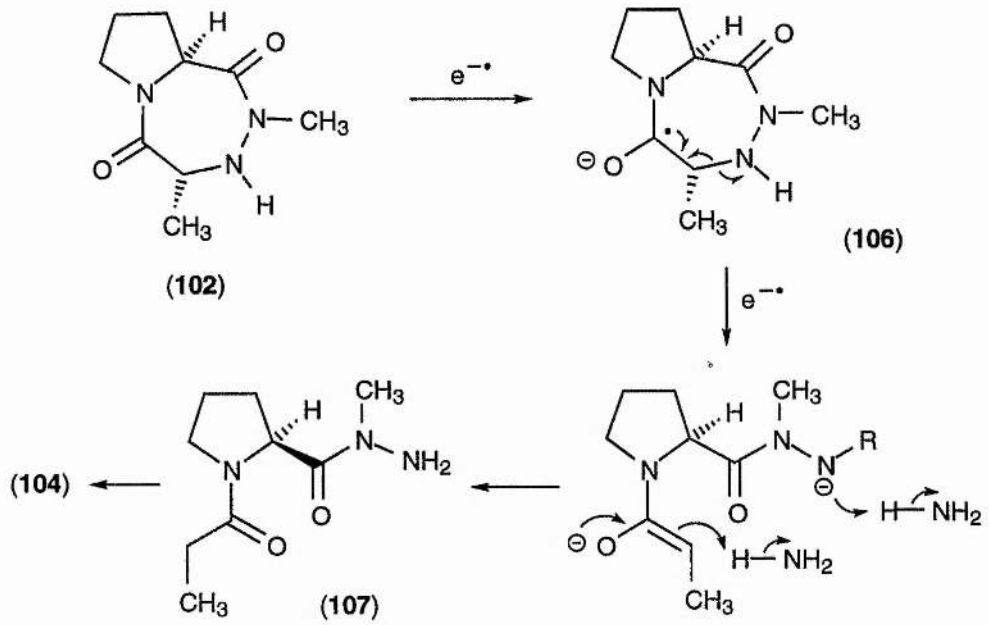
Scheme 2.25: *Alternative synthesis of reduction product (**104**)*

Hesse *et al.* reported a similar reaction for (**105**), an intermediate in the synthesis of the spermidine alkaloid *N*-Acetyl-*N'*-deoxymayfoline.[232] On treatment with sodium in liquid ammonia, a hydrogenolytic C(5)-N(6) fission occurred in preference to the desired N(1)-N(6) cleavage. It was concluded that, under these conditions, C-N bond cleavage will be more favourable than N-N cleavage if the latter is allylic to a double bond. The N-N bond in (**102**) and (**103**) can be thought of as heteroallylic as it is α to a carbonyl group.



(**105**)

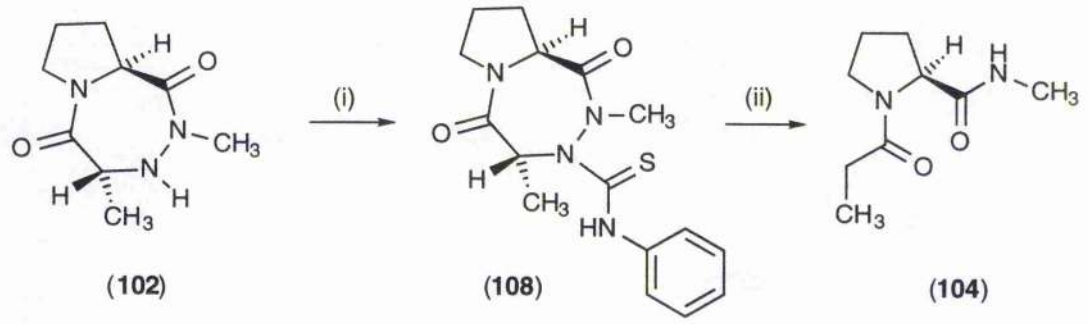
Catalytic hydrogenation of (**102**) and (**103**), using palladium on charcoal, also resulted in C-N bond cleavage and the formation of (**104**). The reactivity of these compounds with sodium in liquid ammonia can be rationalized by the mechanism shown in Scheme 2.26. If the prolyl amide carbonyl acts as an acceptor for a solvated electron produced by the dissolved sodium, a radical anion, (**106**), would be formed where the radical is adjacent to the C-N bond. Homolytic cleavage of the C-N bond would then follow to give the intermediate, (**107**) which would undergo subsequent N-N bond cleavage and loss of ammonia to give (**104**).



Scheme 2.26: *Proposed reduction mechanism*

In the case of the reduction of the bicyclic hydrazides, (**49**) and (**50**), the secondary amide carbonyl was probably the electron accepting site (see p. 76). This would have generated a radical next to the N-N bond and resulted in the observed cleavage of the bond. According to this mechanism, if a radical could be produced which was adjacent to the N-N bond in (**102**) or (**103**), then the bond could be cleaved in a similar manner.

The thoiurea derivative (**108**) was prepared from the reaction of (**102**) with phenyl *iso*-thiocyanate in 76% yield, m.p. 100-3 °C (Scheme 2.27). The thioamide double bond was expected to be a better electron acceptor than either of the amide carbonyls and it was hoped to produce a radical at this site. This radical would be α to both the C-N and N-N bonds and it was thought that selective cleavage of the N-N bond would occur. On treatment of (**108**) with sodium in liquid ammonia, however, the reduction product was again (**104**), emphasizing the ease of cleavage of the C-N bond. Clearly, in order to avoid cleavage of this bond, a radical was required which was exclusively α to the N-N bond.



(102) (108) (104)

Reagents and conditions: (i) PhNCS, CH_2Cl_2, 40 °C, 2 h; (ii) $NH_3(l)$, THF, Na, -60 °C, 30 min.

Scheme 2.27: *Synthesis and reactivity of thiourea derivative (**108**)*

The surprising ease of cleavage of the C-N bonds of (**102**) and (**103**) in dissolving metal reductions may be partly due to the orientation of the acceptor carbonyl relative to the C-N bond. Computer modelling studies on the bicyclic compound (**102**) show that the C-N bond is almost orthogonal to the adjacent carbonyl group (Fig. 2.15). This means that the orbital containing the radical would be lined up with the C-N bond in a manner which would facilitate subsequent homolytic cleavage of the bond.

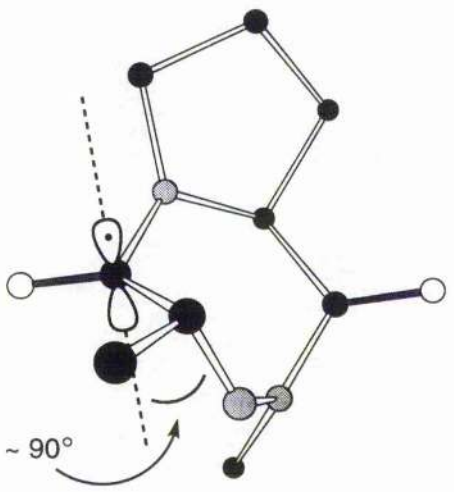
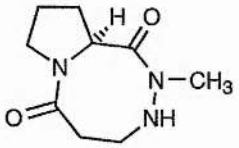


Figure 2.15: *Geometry of C-N bond of (**102**)*

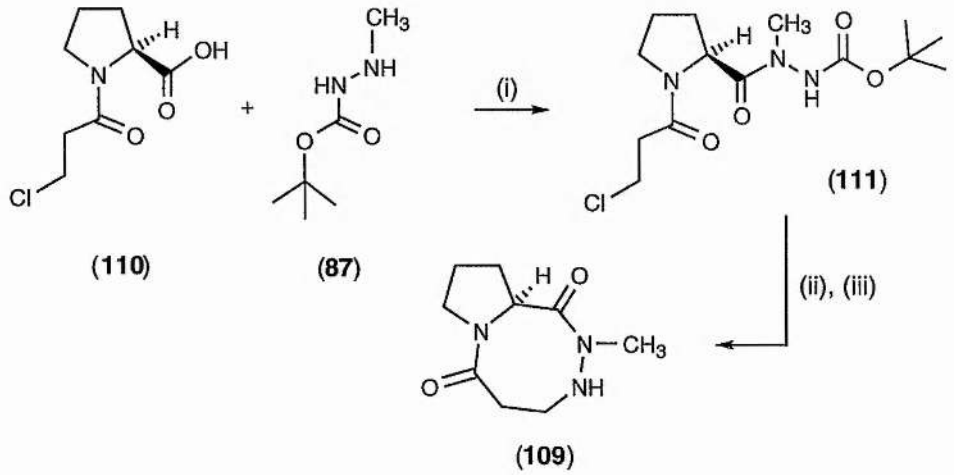
2.5.7 Synthesis and reactivity of an *N*-methyl *cis* βAla-Pro mimetic

The 5,8 bicyclic compound, (**109**), is a *cis* βAla-Pro mimetic. Since the hydrazine C-N bond in this case is β to the nearest carbonyl, it should not be possible for it to cleave *via* the mechanism outlined in Scheme 2.26. In the absence of this competing reaction, it was hoped that the dissolving metal reduction of (**109**) would result in the selective cleavage of the N-N bond.



(**109**)

To synthesize (**109**), *N*-3-chloropropanoyl-(2*S*)-proline (**110**) was coupled to the methylhydrazine derivative (**87**) using mixed anhydride methodology to produce the prolyl hydrazide (**111**) in 88% yield, m.p. 108-11 °C (Scheme 2.28). Removal of the ^{t}BOC group, followed by the subsequent cyclization of the resulting hygroscopic salt using 1 mol dm^{-3} NaOH gave the novel 5,8-bicyclic pyrrolo-triazocine compound (**109**), m.p. 129-32 °C. Although the formation of the eight-membered rings is generally considered to be less favoured than the formation of seven-membered rings, the cyclization to produce (**109**) was very rapid, as judged by TLC.

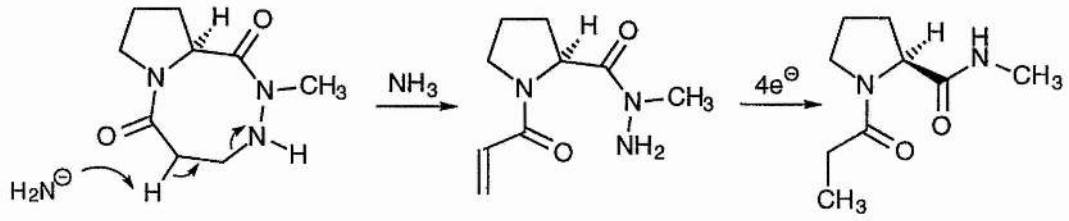


(**110**) (**87**) (**111**) (**109**)

Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF, -15 °C, 1.5 h, 88%; (ii) HCl(g), EtOAc, 0 °C, 15 min; (iii) 1 mol dm^{-3} NaOH, MeOH, 5 min, 60%.

Scheme 2.28: *Synthesis of an* N*-methyl* cis *βAla-Pro mimetic*

It was hoped that the N-N bond of (**109**) could be selectively cleaved to give an open chain dipeptide, but reduction of (**109**) with sodium in liquid ammonia gave *N*-propionyl-(2*S*)-proline methylamide (**104**) in almost quantitative yield. It seems likely that in the presence of a strong base (such as sodamide), compound (**109**) underwent a *retro*-Michael reaction (Scheme 2.29) to give the acryloyl proline intermediate (**112**). The double bond of (**112**) would be rapidly reduced under the conditions used, and the terminal hydrazide cleaved, to give the final product, (**104**).



(**109**) (**112**) (**104**)

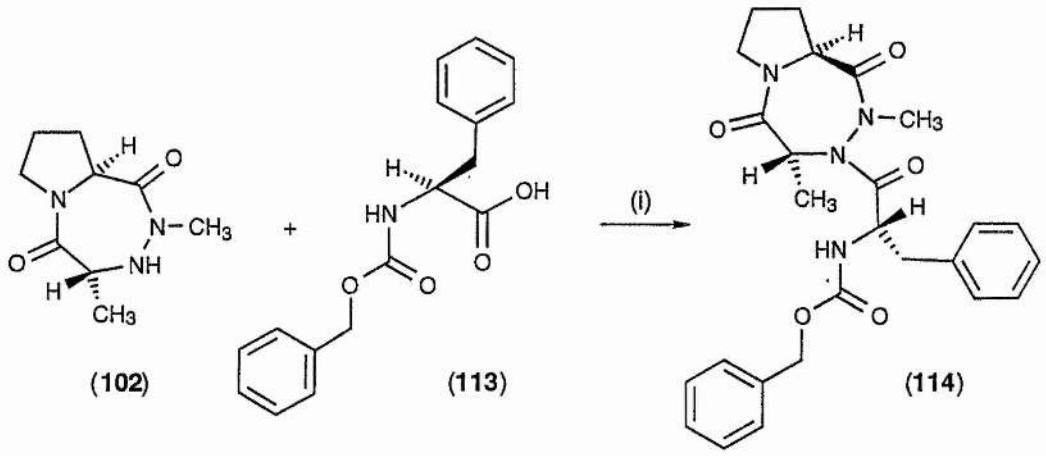
Scheme 2.29: *Proposed mechanism of reduction of (109)*

Although the reduction of the 5,8-bicyclic compound (**109**) did not give the desired product, it has been shown that *cis* X-Pro mimetics where X is a β-amino acid, can be easily prepared using the synthetic protocols outlined in section 2.5.5 (p. 93) for α-amino acids.

2.5.8 Extension of *cis* X-Pro mimetics

It was important that the new bicyclic *cis* X-Pro mimetics could be extended at both the amino and carboxy termini to provide longer peptide mimetics. In particular, it was hoped to incorporate this unit into the FMDV 2A sequence in order to study the structural effects.

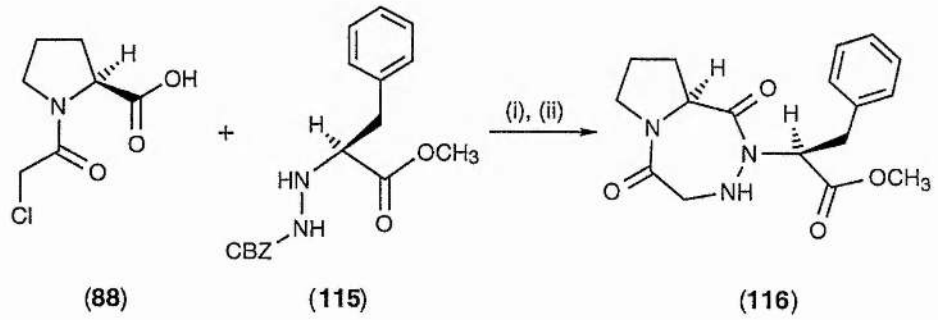
In order to assess the possibility that the triazepine compounds might be extended at the N-terminal to provide more elaborate peptide mimetics, compound (**102**) was reacted with *N*-benzyloxycarbonyl-(2*S*)-phenylalanine (**113**) using the mixed anhydride coupling method (Scheme 2.30). The product (**114**), a protected *cis* Phe-Gly-Pro mimetic was obtained in excellent yield (82%), m.p. 82-5 °C. This demonstrated that N-terminal extensions with acyl group activated peptides should be straightforward.



Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF, -15 °C, 1.5 h, 82%.

Scheme 2.30: *N-terminal extension of a* cis *Ala-Pro mimetic*

Extension at the proline carboxamide terminal is more difficult, requiring the synthesis of unsymmetrical alkylhydrazines protected on the primary amino group, such as the methylhydrazine derivative (**87**). Nevertheless, the successful synthesis of the *cis* Gly-Pro-Phe methyl ester mimetic (**116**) has been carried out by co-workers[233] using *N*-chloroacetyl-(2*S*)-proline (**88**) and a protected form of hydrazino-(2*S*)-phenylalanine (**115**).



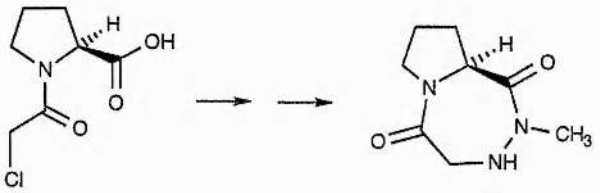
Reagents and conditions: (i) $SOCl_2$, DMAP, rt., 12 h; (ii) HBr, AcOH, 3h; NMM.

Scheme 2.31: *Synthesis of a* cis *Gly-Pro-Phe mimetic*

2.6 Synthesis and reactivity of glyoxyl-proline

2.6.1 Introduction

The successful preparation of the *cis* Gly-Pro bicyclic compound, (**91**) was achieved by reacting the α-chloro acid (**88**) with a protected form of methylhydrazine. It was important that the compound had the correct regiochemistry. This would be the case if the primary group of methylhydrazine had attacked at the halogenated carbon. Substitution of the α-chloro group with an aldehyde should facilitate preferential attack of a primary amine at this site.



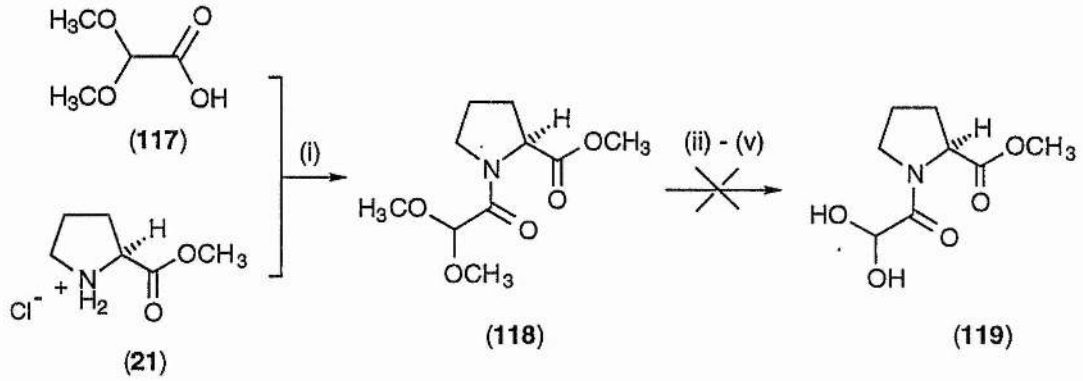
(**88**) (**91**)

As part of an earlier study, we had wished to synthesize the *N*-glyoxyl-prolyl compound (**119**), which should react with methylhydrazine to give an alternative synthesis of the *cis* Gly-Pro bicyclic compound, (**91**).

2.6.2 Attempted syntheses of *N*-glyoxyl proline

The aldehyde function is particularly unstable when situated α to an amide group and would almost certainly exist in its hydrated form, as in (**119**). The best strategy towards the synthesis of this compound would be to prepare the glyoxyl proline in a protected form, followed by subsequent deprotection. The dimethyl acetal of glyoxylic acid, (**117**), was coupled to proline methyl ester (**21**) using the mixed anhydride method to give dimethoxyacetyl-(2*S*)-proline methyl ester, (**118**) in 91% yield as an oil (Scheme 2.32).

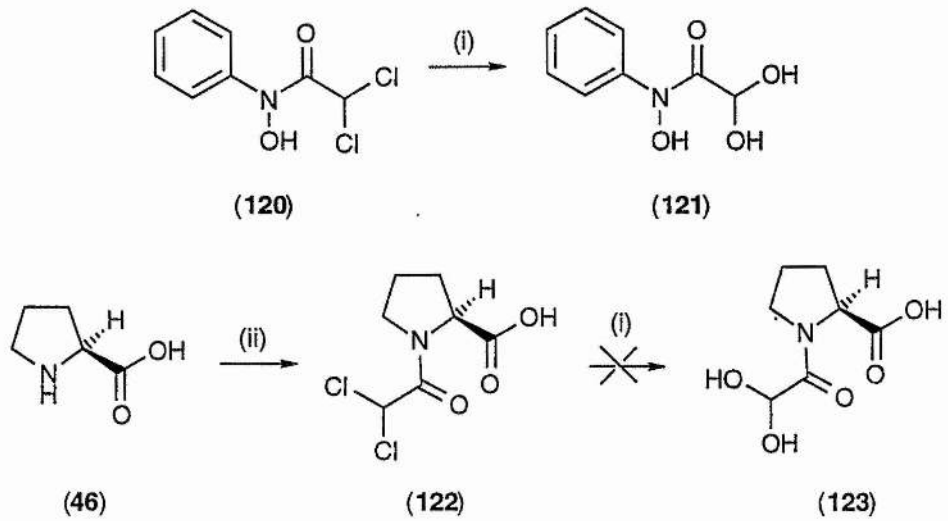
The conversion of acetals into aldehydes is normally accomplished by acid-catalyzed hydrolysis,[234] but the protected glyoxyl proline (**118**) proved to be completely unreactive under these conditions. ^{1}H NMR showed only the presence of unreacted starting material after treatment with HCl (2 mol dm^{-3}) for 12 h. Various other deprotection methods, including use of the electrophilic reagent, trimethylsilyl iodide,[235] also proved unsuccessful.



Reagents and conditions: (i) NMM, *iso*-butyl chloroformate, THF/ DMF, -15 °C, 1.5 h, 91%; (ii) HCl (2 mol dm^{-3}), MeOH/ H_2O (1:1), 20 °C, 12 h; (iii) CF_3CO_2H/ CH_2Cl_2 (1:1), 20 °C, 2 h; (iv) *p*-TSA, acetone, reflux, 12 h; (v) Me_3SiI, CH_2Cl_2, 20 °C, 15 min.

Scheme 2.32: *Attempted synthesis of* N-*glyoxyl-proline methyl ester*

Another masked form of the aldehyde is the *N*-α,α-dichloroacetyl proline derivative, (**122**). Smissman *et al.* reported the synthesis of *N*-phenylglyoxylhydroxamic acid monohydrate, (**121**) from the base-catalyzed hydrolysis of *N*-phenyl-α,α-dichloroaceto-hydroxamic acid, (**120**) (Scheme 2.33).[236] Dichloroacetyl-(2*S*)-proline (**122**) was easily prepared from proline (**46**) in 63% yield, m.p. 133-5 °C. This compound, however, proved to be completely stable to hydrolysis, even in strongly basic solution and at elavated temperatures.



(120) (121)

(46) (122) (123)

Reagents and conditions: (i) NaOH (2 mol dm^{-3}), 2 h; (ii) $Cl_2CHCOCl$, Et_2O, $NaHCO_3$, 0 °C, 30 min.

Scheme 2.33: *Hydrolysis of α,α-dichloroacetyl compounds*

It seemed clear that the N-CO-CHO fragment was very unstable and that its protected forms were extremely difficult to unblock. As formation of compound (**123**) proved difficult, we decided to synthesize (**127**) through the ozonolysis of *N*-crotonyl-proline methyl ester (**126**) (Scheme 2.34).

2.6.3 Synthesis and reactivity of *N*-glyoxyl proline

N-Crotonyl-proline (**125**), was prepared by coupling *trans* crotonyl chloride (**124**) to proline (**46**) in 64% yield, m.p. 161-3 °C. The ^1H and ^{13}C NMR spectra of (**125**) showed that it contained an unusually high percentage of the *cis* isomer in solution (37%). A single crystal X-ray diffraction study of (**125**) was carried out in order to establish its conformation in the solid state. In the crystal lattice, the prolyl amide bond was found to be in the *trans* conformation (Fig. 2.16).

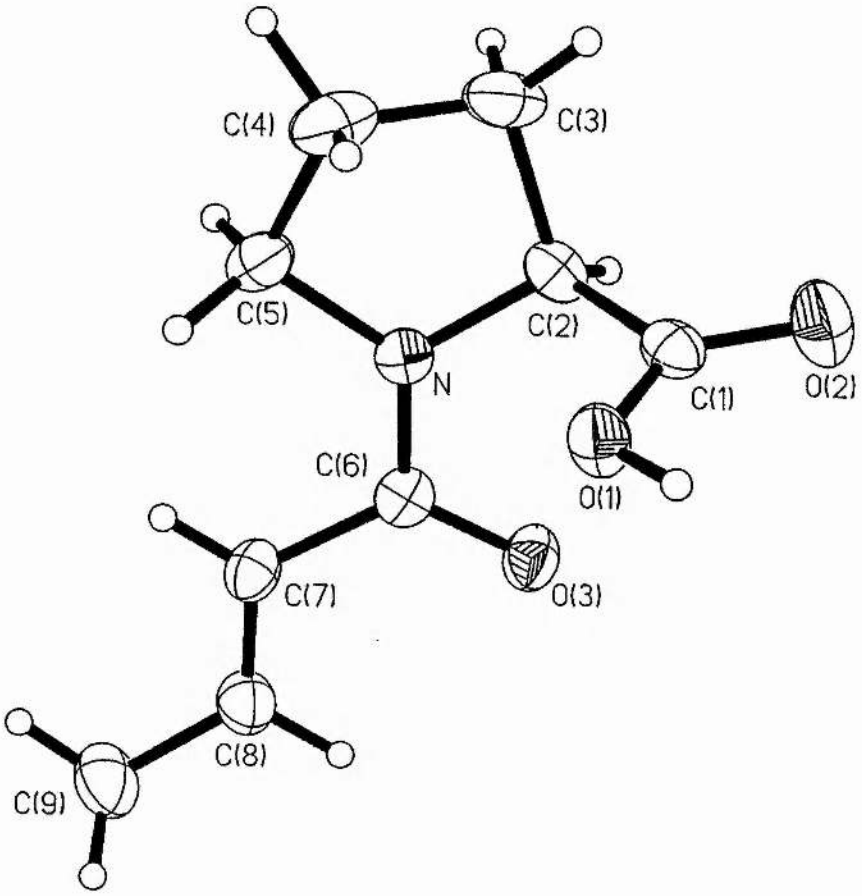
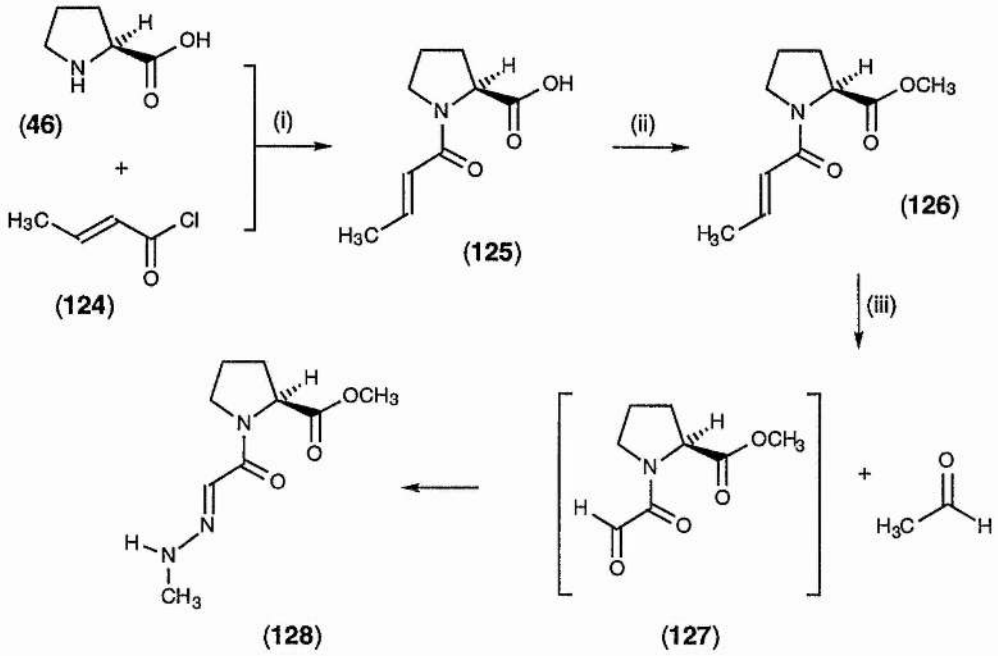


Figure 2.16: *X-ray structure of (125)*

N-Crotonyl-proline (**125**), was then esterified under non-acidic conditions, using the methanol adduct of BF_3, to give *N*-crotonyl-proline methyl ester (**126**) (Scheme 2.34).[237] Ozonolysis of (**126**) in CH_2Cl_2 was complete after approximately one hour, as judged by TLC and dimethylsulfide was used to break down the intermediate ozonide. Due to its inherent instability, the aldehyde (**127**) was directly treated with methylhydrazine *in situ*.

The NMR spectra of the resultant product were extremely complex. It appeared that two major species and two minor species were present, which was suggestive of conformational isomerism. Mass spectral data confirmed that the structure of the product was that of the hydrazone (**128**) ($[M + H]^+ = 214$). Molecular modelling studies suggested that (**128**) could exist in four relatively stable conformations, due to

cis/trans isomerism of both the imine bond and the amide bond (Fig. 2.17). This would explain the presence of four species in the NMR spectra.



Reagents and conditions: (i) NaOH, 0 °C, 1 h; HCl, 64%; (ii) BF_3.MeOH, reflux, 2 h; $NaHCO_3$, 73%; (iii) O_3, CH_2Cl_2, -78 °C, 1 h; Me_2S, $NHCH_3NH_2$, rt., 8 h, 94%.

Scheme 2.34: *Synthesis and reactivity of* N-*glyoxyl-proline methyl ester*

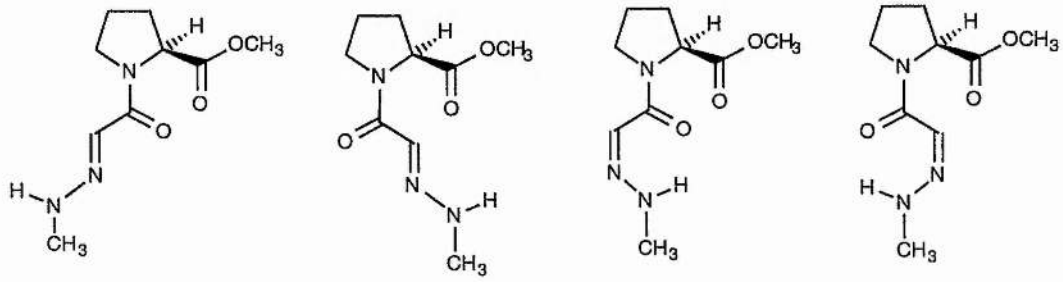
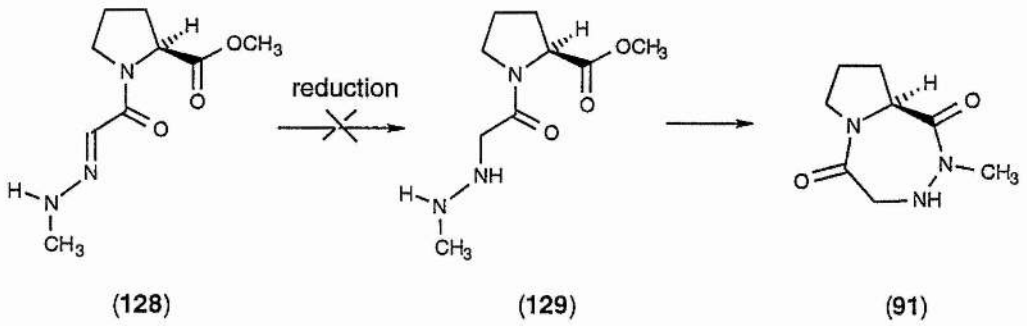


Figure 2.17: *Possible conformers of (128)*

The hydrazone, (**128**) was structurally rather rigid and hence unable to undergo intramolecular cyclization. It was hoped that (**128**) could be reduced to give the more flexible hydrazine compound, (**129**). The terminal amine group of (**129**) would then be able to attain the correct conformation for cyclization to occur (Scheme 2.35). Hydrazones are normally reduced to amines with the use of reducing agents such as $NaBH_4$, $LiAlH_4$ or sodium in ethanol. Compound (**128**), however, could not be reduced by any of these methods, as judged by NMR. It can be concluded that the hydrazone is rather unreactive due to the presence of the adjacent amide group.



Scheme 2.35: *Proposed synthesis of (91)*

Although this line of investigation did not provide a viable synthetic route to the bicyclic compound (**91**), we were able to synthesize successfully an *N*-glyoxyl-proline compound *via* ozonolysis (Scheme 2.34). It should be noted that due to its inherent instability, the aldehyde should be generated *in situ* for immediate reaction with an appropriate nucleophile.

2.7 Conclusions and future work

A successful route has been developed for the synthesis of *cis* constrained, prolyl dipeptides. These bicyclic compounds could be incorporated into larger peptides and used as β VI turn mimetics (see section 1.5, p. 44). The reactivity of these compounds has been investigated with a view to their conversion into natural *cis* X-Pro peptides *via* low temperature, reductive cleavage. Although dissolving metal reduction has not resulted in the formation of the desired open chain dipeptides, information has been gained on the reactivity of the bicyclic compounds.

In the future, it is hoped to prepare bicyclic *cis* X-Pro mimetics that are substituted on the amide nitrogen with a radical-producing group. If a radical could be generated in a position exclusively α to the N-N bond, it is hoped that this might result in selective cleavage of the bond. Oligopeptides containing the FMDV 2A sequence could be synthesized containing a *cis* Gly-Pro mimetic unit and it is hoped that this unit could be converted to the natural peptide under mild conditions. This peptide might then be able to attain the correct conformation to exhibit self-cleavage properties, as observed in the replication process of the virus itself.

Studies on oligopeptides containing the FMDV 2A sequence have indicated the presence of α-helical structure in the N-terminal region of these peptides in non-polar media. In the future it is planned to synthesize the oligopeptides with helix-capping groups at the amino terminal.[238] It is hoped that the presence of these groups will result in the formation of a more stable α-helix in the peptides, which might be a structural prerequisite for auto-hydrolysis to occur.

The influence of prolyl *cis/trans* isomerism on the structure of polypeptides is poorly understood. The mechanisms of the isomerization process in free solution and at the active site of peptidyl-prolyl *cis-trans* isomerases (*e.g.* cyclophilin) are also not known. A synthetic approach to the preparation of *cis* X-Pro peptides, such as the one reported here would facilitate a comprehensive kinetic and mechanistic examination of the isomerization process.

CHAPTER 3: EXPERIMENTAL

Experimental Procedure:

Elemental microanalyses were performed in the departmental microanalytical laboratory.

NMR spectra were recorded on a Brüker AM-300 (300 MHz; f.t. ^{1}H-NMR, and 74.76 MHz; ^{13}C-NMR), Varian gemini 200 (200 MHz; f.t. ^{1}H-NMR and 50.31 MHz; ^{13}C-NMR) or by the SERC service at Warwick using a Brüker AM-400 (400 MHz; f.t. ^{1}H-NMR, and 100 MHz; ^{13}C-NMR) spectrometers. ^{1}H-NMR and ^{13}C-NMR spectra are described in parts per million downfield from TMS and are reported consecutively as position (δ_H or δ_C), relative integral, multiplicity (s -singlet, d -doublet, t -triplet, q- quartet, dd -doublet of doublets, sep -septet, m -multiplet and br -broad), coupling constant (Hz) and assignment (numbering according to the IUPAC nomenclature for the compound). ^{1}H-NMR were referenced internally on ^{2}HOH (4.68 ppm), $CHCl_3$ (7.27 ppm) or DMSO (2.47 ppm). ^{13}C-NMR were referenced on CH_3OH (49.3 ppm), $C^{2}HCl_3$ (77.5 ppm), or DMSO (39.7 ppm).

Pyrrolidine ring carbons and hydrogens are assigned in NMR spectra as α, β, γ and δ, going anticlockwise from the ring nitrogen, according to normal convention. Where more than one conformational isomer can be seen in the NMR spectrum due to the presence of a tertiary amide bond, these are assigned as *c* (*cis*) or *t* (*trans*), according to the isomeric state of the amide bond. When two such tertiary amide bonds are contained in one compound and four conformational isomers are present, these are assigned as *tt* (*trans, trans*), *tc* (*trans, cis*), *ct* (*cis, trans*) and *cc* (*cis, cis*). The first assignment refers to the amide bond on the left hand side of the molecule as it is pictured diagramatically in the preceding chapter. Where two sets of peaks arise in NMR spectra due to different conformations of a constrained seven-membered ring, these are assigned as *A* and *B*, with *A* being the major isomer.

IR spectra were recorded on a Perkin-Elmer 1710 f.t. IR spectrometer. The samples were prepared as Nujol mulls, solutions in chloroform or thin films between sodium chloride discs. The frequencies (υ) as absorption maxima are given in wavenumbers (cm^{-1}) relative to a polystyrene standard. Mass spectra and accurate mass measurements were recorded on a VG 70-250 SE, a Kratos MS-50 or by the SERC service at Swansea using a VG AZB-E. Fast atom bombardment spectra were recorded using glycerol as a matrix. Major fragments were given as percentages of the base peak intensity (100%). UV spectra were recorded on Pye-Unicam SP8-500 or SP8-100 spectro-photometers.

Flash chromatography was performed according to the method of Still *et al.*[65] using Sorbsil C 60 (40-60 μm mesh) silica gel. Analytical thin layer chromatography was carried out on 0.25 mm precoated silica gel plates (Macherey-Nagel SIL g/UV254) and compounds were visualised using UV fluorescence, iodine vapour, ethanolic phosphomolybdic acid or ninhydrin.

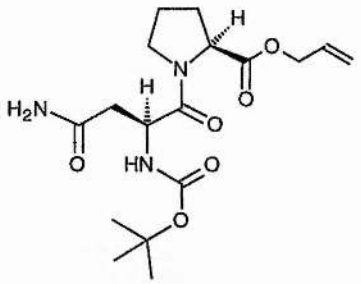
Melting points were taken on an Electrothermal melting point apparatus and are uncorrected. Optical rotations were measured at 22 °C on an Optical Activity AA-1000 polarimeter using 10 or 20 cm path length cells.

The solvents used were either distilled or of analar quality and light petroleum ether refers to that portion boiling between 40 and 60 °C. Solvents were dried according to literature procedures. Ethanol and methanol were dried using magnesium turnings. DMF, *iso*propanol, *iso*pentanol, toluene, CH_2Cl_2, acetonitrile, di*iso*propylamine, triethylamine and pyridine were distilled over CaH_2. THF and diethylether were dried over sodium/ benzophenone and distilled under nitrogen. Thionyl chloride was distilled over sulphur and the initial fractions were always discarded. *N*-Methylmorpholine was distilled over ninhydrin.

General procedure for dissolving metal reductions

To liquid ammonia (40 cm^3) at -60 °C under an atmosphere of dry nitrogen was added a solution of the starting material (10 mmol) in dry THF (10 cm^3). If the starting material was found to be insoluble in THF, it was added directly to the liquid ammonia. The resultant solution was stirred vigorously and sodium (approx 0.8 g, 30 mmol) was added in small pieces until a homogenous dark blue colour was obtained. This blue colour was maintained for 15 min, after which time the solution cleared and ammonium chloride (1.60 g, 30 mmol) was added. The reaction mixture was allowed to warm to room temperature, methanol (20 cm^3) was added, and the mixture was left stirring for 1 h. The solvents were removed under reduced pressure and the resultant white solids* were dissolved in water (25 cm^3) and extracted with ethyl acetate (3 x 25 cm^3). The combined organic fractions were dried ($MgSO_4$) and the solvent removed under reduced pressure to give the reduced product. In the case of polar products, the white solids* were extracted into hot ethanol (2 x 25 cm^3), the solvent removed under reduced pressure and the resultant reduced product recrystallized from methanol.

***N-tert*-Butoxycarbonyl-(2*S*)-asparagyl-(2*S*)-proline allyl ester (16)**



To a solution of *N-tert*-butoxycarbonyl-(2*S*)-asparagine (**14**) (1.16 g, 5 mmol) in dry THF (250 cm^3) was added *N*-methylmorpholine (0.56 cm^3, 5 mmol) and the solution cooled to -15 °C. *iso*-Butyl chloroformate (0.68 cm^3, 5 mmol) was added with stirring and the solution left stirring at -15 °C for 2 min. A solution of (2*S*)-proline allyl ester *p*-toluene sulphonate (**15**) (1.64 g, 5 mmol) and *N*-methylmorpholine (0.56 cm^3, 5 mmol) in dry THF (20 cm^3) was then added The reaction mixture was allowed to warm to room temperature and left stirring for 2 h. The hydrochloride salts were filtered off and the solvents removed under reduced pressure. The resultant yellow oil was dissolved in CH_2Cl_2 (50 cm^3) and washed with 0.5 mol dm^{-3} HCl (2 x 15 cm^3) and 5% sodium bicarbonate solution (2 x 15 cm^3). The organic phase was dried ($MgSO_4$) and the solvent removed under reduced pressure. The resultant white solid was recrystallised from diethyl ether/ light petroleum to give the product as colourless crystals (1.17 g, 63%), m.p. 109-11 °C; (Found: C, 55.4; H, 7.65; N, 11.4. Calc. for $C_{17}H_{27}N_3O_6$: C, 55.25 H, 7.4; N, 11.4%); *m/z* (Found: $[M + H]^+$, 370.1978. Calc. for $C_{17}H_{28}N_3O_6$: 370.1978); $[\alpha]_D^{22}$ -78.0 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3458 & 3305 (NH), 1743 (ester CO), 1694 (carbamate CO), 1674 (tertiary amide CO), 1644 (primary amide CO) and 1521 (C=C); δ_H (300 MHz; C^2HCl_3) 1.41 (9H, s, $(CH_3)_3$), 2.02 (3H, m, $\frac{1}{2}\beta CH_2$ and γCH_2), 2.22 (1H, m, $\frac{1}{2}\beta CH_2$), 2.56 (1H, dd, J_1 14.5, J_2 5.8, $\frac{1}{2}\underline{C}H_2CONH_2$), 2.68 (1H, dd, J_1 14.5, J_2 5.6, $\frac{1}{2}C\underline{H}_2CONH_2$), 3.79 (2H, m, δCH_2), 4.54 (1H, dd, J_1 5.7, J_2 9.8, αCH), 4.58 (2H, d, *J* 5.3, CO_2CH_2), 4.82 (1H, m, COC<u>H</u>NH), 5.24 (1H, d, *J* 10.4, $=C\underline{H}_{cis}H$), 5.31 (1H, d, *J* 17.3, $=C\underline{H}_{trans}H$), 5.68 (1H, s, $\frac{1}{2}CONH_2$), 5.88 (1H, m, C<u>H</u>=CH_2), 6.02 (1H, d, *J* 8.0, CHN<u>H</u>) and 6.70 (1H, s, $\frac{1}{2}CONH_2$); δ_C (74.76 MHz; C^2HCl_3) 22.39 (*c*, γCH_2), 25.25 (*t*, γCH_2), 28.64 (*t* & *c*, $C(\underline{C}H_3)_3$), 29.28 (*t*, βCH_2), 31.44 (*c*, βCH_2), 38.04 (*t*, $\underline{C}H_2CONH_2$), 39.30 (*c*, $\underline{C}H_2CONH_2$), 46.90 (*c*, δCH_2), 47.45 (*t*, δCH_2), 49.52 (*c*, <u>C</u>HNHCO), 49.85 (*t*, <u>C</u>HNHCO), 59.47 (*t*, αCH), 59.94 (*c*, αCH), 65.97 (*t*, $CO_2\underline{C}H_2$), 66.37 (*c*, $CO_2\underline{C}H_2$), 79.86 (*c*, $\underline{C}(CH_3)_3$), 79.98 (*t*, $\underline{C}(CH_3)_3$), 118.80 (*t*, CH=$\underline{C}H_2$), 118.91 (*c*, CH=$\underline{C}H_2$), 132.10 (*t*, <u>C</u>H=CH_2), 132.20 (*c*, <u>C</u>H=CH_2), 155.15 (*c*, $CO_2{}^tBu$), 155.85 (*t*, $CO_2{}^tBu$), 171.14 (*c*, $CONH_2$), 171.49 (*t*, $CONH_2$), 171.97 (*t* & *c*,

COCHNH), 172.72 (*t*, $\underline{C}O_2CH_2$) and 173.32 (*c*, $\underline{C}O_2CH_2$); *m/z* (CI) 370 ([*M* + H]$^+$, 44%), 314 (52, [*M* - $C(CH_3)_3$ + 2H]$^+$), 296 (30, [*M* - $OC(CH_3)_3$]$^+$), 270 (40 [*M* - $CO_2C(CH_3)_3$ + 2H]$^+$), 255 (100, [*M* - $NHCO_2C(CH_3)_3$ + 2H]$^+$), 212 (100, [*M* - $NHCO_2C(CH_3)_3$ - C_3H_5]$^+$) and 70 (45, [C_4H_8N]$^+$).

N-tert-Butoxycarbonyl-(2*S*)-asparagyl-(2*S*)-proline benzyl ester (18)

This compound was prepared in a manner identical with that for *N-tert*-butoxycarbonyl-(2*S*)-asparagyl-(2*S*)-proline allyl ester (**16**), using *N-tert*-butoxycarbonyl-(2*S*)-asparagine (**14**) (1.16 g, 5 mmol) and (2S)-proline benzyl ester *p*-toluene sulphonate (**17**) (1.89 g, 5 mmol). The crude compound was recrystallised from CH_2Cl_2/ light petroleum to give a white solid (1.67 g, 79%), m.p. 111-13 °C (lit.,[190] 115-16 °C); *m/z* (Found: [*M* + H]$^+$, 420.2140. $C_{21}H_{30}N_3O_6$ requires 420.2135); $[\alpha]_D^{22}$ -56.4 (*c* 1.0 in MeOH) {lit.,[190] -53.8 (*c* 1.0 in MeOH)}; υ_{max} (Nujol)/cm^{-1} 3393 & 3246 (NH), 1743 (ester CO), 1697 (carbamate CO), 1672 (tertiary amide CO), 1643 (primary amide CO) and 1045 & 1026 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.42 (9H, s, $C(CH_3)_3$), 1.98-2.09 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.11-2.32 (1H, m, $\frac{1}{2}\beta CH_2$), 2.56 (2H, m, $COCH_2$), 3.76 (2H, m, δCH_2), 4.57 (1H, dd, J_1 4.0, J_2 8.2, αCH), 4.88 (1H, m, C<u>H</u>NH), 5.12 (2H, d, *J* 5.0, CH_2Ph), 5.40 (1H, s, $\frac{1}{2}CONH_2$), 5.91 (1H, d, *J* 8.2, CHN<u>H</u>), 6.48 (1H, s, $\frac{1}{2}CONH_2$) and 7.33 (5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 25.31 (γCH_2), 28.82 ($C(\underline{C}H_3)_3$), 29.39 (βCH_2), 38.30 ($CO\underline{C}H_2$), 47.62 (δCH_2), 50.00 ($\underline{C}H_2CONH_2$), 59.67 (αCH), 67.31 (<u>C</u>HNH), 80.28 ($\underline{C}(CH_3)_3$), 128.56 (*ortho* aromatic), 128.82 (*para* aromatic), 129.06 (*meta* aromatic), 135.99 (quat. aromatic), 156.00 ($CO_2{}^tBu$), 171.61 (<u>C</u>ONHCH), 172.28 ($CONH_2$) and 172.61 ($\underline{C}O_2CH_2Ph$); *m/z* (FAB) 420 ([*M* + H]$^+$, 36%), 364 (39, [*M* - $C(CH_3)_3$ + 2H]$^+$), 320 (43, [*M* - $CO_2C(CH_3)_3$ + 2H]$^+$), 287 (40, [*M* - CO_2CH_2Ph + 3H]$^+$), 70 (22, [C_4H_8N]$^+$) and 57 (100, [$C(CH_3)_3$]$^+$).

N-tert-Butoxycarbonyl-(2*S*)-asparagyl-(2*S*)-proline (19)

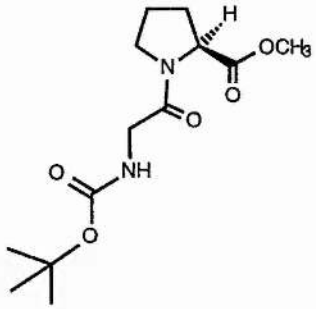
To a solution of *N-tert*-butoxycarbonyl-(2*S*)-asparagyl-(2*S*)-proline benzyl ester (**18**) (2.10 g, 5 mmol) in methanol (100 cm^3) was added 10% palladium on activated charcoal (100 mg) and the vessel flushed with hydrogen gas. The resultant suspension was stirred vigorously under an atmosphere of hydrogen for 12 h. The mixture was then filtered through a prewashed celite pad and the solvent removed under reduced pressure to give the product as a white solid (1.55 g, 94%),

m.p. 110-15 °C (lit.,[190] 108-16 °C); *m/z* (Found: $[M + H]^+$, 330.1659. $C_{14}H_{24}N_3O_6$ requires 330.1665); $[\alpha]_D^{22}$ -46.8 (*c* 1.0 in MeOH) {lit.,[190] -50.0 (*c* 1.0 in MeOH)}; υ_{max} (Nujol)/cm^{-1} 3405 & 3198 (NH), 1717 (acid CO), 1692 (carbamate CO), 1669 (tertiary amide CO) and 1645 (primary amide CO); δ_H (200 MHz; C^2HCl_3) 1.39 (9H, s, $C(CH_3)_3$), 1.78-2.36 (4H, m, βCH_2 and γCH_2), 2.64 (2H, m, $COCH_2$), 3.83 (2H, m, δCH_2), 4.39 (1H, m, αCH), 4.82 (1H, m, C<u>H</u>NH), 6.24 (1H, d, *J* 7.8, CHN<u>H</u>), 6.73 (1H, s, $\frac{1}{2}CONH_2$), 7.22 (1H, s, $\frac{1}{2}CONH_2$) and 8.12 (1H, br, CO_2H); δ_C (50.31 MHz; C^2HCl_3) 24.51 (γCH_2), 28.00 ($C(\underline{C}H_3)_3$), 28.63 (βCH_2), 36.98 ($CO\underline{C}H_2$), 47.02 (δCH_2), 48.96 (<u>C</u>HNH), 59.27 (αCH), 79.55 ($\underline{C}(CH_3)_3$), 155.24 ($CO_2{}^tBu$), 170.96 (<u>C</u>ONHCH), 173.62 ($CONH_2$) and 174.12 (CO_2H); *m/z* (CI) 330 ($[M + H]^+$, 23%), 274 (51, $[M - C(CH_3)_3 + 2H]^+$), 230 (33, $[M - CO_2C(CH_3)_3 + 2H]^+$), 70 (72, $[C_4H_8N]^+$) and 57 (100, $[C(CH_3)_3]^+$).

(2*S*)-Proline methyl ester hydrochloride (21)

Thionyl chloride (1.6 cm^3, 22 mmol) was added dropwise with stirring to dry methanol (30 cm^3) at 0 °C. The resultant solution was stirred at 0 °C for 5 min before addition of (2*S*)-proline (**47**) (2.30 g, 20 mmol). When dissolution was complete, the solution was refluxed for 1 h, and after cooling, methanol and excess thionyl chloride were removed under reduced pressure to give the product as a clear oil. After drying over P_2O_5 for 48 h the hydrochloride was obtained as colourless crystals (3.31 g, 100%), m.p. 68-70 °C (lit.,[239] 69-71 °C); *m/z* (Found: $[M + H]^+$, 130.1663. $C_6H_{12}NO_2$ requires 130.1663); $[\alpha]_D^{22}$ -14.8 (*c* 1.0 in H_2O) {lit.,[239] -15.5 (*c* 1.0 in H_2O)}; υ_{max} (Nujol)/cm^{-1} 1746 (CO) and 1240 (C-O); δ_H (200 MHz; d_6-DMSO) 1.80-2.08 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.12-2.34 (1H, m, $\frac{1}{2}\beta CH_2$), 3.21 (2H, m, δCH_2), 3.72, (3H, s, OCH_3), 4.34 (1H, m, αCH), 9.30 (1H, s, NH) and 10.49 (1H, s, NH); δ_C (50.31 MHz; d_6-DMSO) 23.27 (γCH_2), 27.84 (βCH_2), 45.32 (δCH_2), 53.21 (OCH_3), 59.61 (αCH) and 169.26 (CO); *m/z* (EI) 130 ($[M + H]^+$, 24%), 99 (62, $[M - OCH_3]^+$) and 70 (100, $[C_4H_8N]^+$).

***N*-*tert*-Butoxycarbonyl-glycyl-(2*S*)-proline methyl ester (22)**



To a stirred solution of *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry THF (25 cm^3) was added *N*-*tert*-butoxycarbonyl-glycine (**20**) (1.75 g, 10 mmol) and the solution cooled to -15 °C. *iso*-Butyl chloroformate (1.36 cm^3, 10 mmol) was added and the resulting suspension stirred at -15 °C for 2 min. A mixture of (2*S*)-proline methyl ester hydrochloride (**21**) (1.66 g, 10 mmol) and *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry DMF (5 cm^3) was then added. The reaction mixture was allowed to warm to room temperature and left to stir for 1 h. The hydrochloride salts were filtered and the solvents removed under reduced pressure. The resultant clear oil was dissolved in CH_2Cl_2 (50 cm^3), washed with 0.5 mol dm^{-3} HCl (2 x 25 cm^3) and 5% sodium bicarbonate solution (2 x 25 cm^3), dried ($MgSO_4$) and the solvent removed under reduced pressure. The crude product was recrystallised from diethyl ether/ light petroleum to give colourless crystals (2.54 g, 89%), m.p. 66-8 °C; (Found: C, 54.8; H, 8.15; N, 9.8. Calc. for $C_{13}H_{22}N_2O_5$: C, 54.6; H, 7.75; N, 9.8%); *m/z* (Found: $[M + H]^+$, 287.1610. Calc. for $C_{13}H_{23}N_2O_5$: 287.1610); $[\alpha]_D^{22}$ -34.7 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3404 (NH), 1748 (ester CO), 1686 (carbamate CO) and 1673 (amide CO); δ_H (400 MHz; C^2HCl_3) 1.39 (*t* & *c*, 9H, s, $(CH_3)_3$), 1.76-2.27 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.38-3.45 (*t*, 1H, m, $\frac{1}{2}\delta CH_2$), 3.48-3.65 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 3.67 (*t*, 3H, s, OCH_3), 3.71 (*c*, 3H, s, OCH_3), 3.86 (*t*, 1H, dd, J_1 4.0, J_2 17.5, $\frac{1}{2}COCH_2$), 3.93 (*c*, 1H, dd, J_1 4.6, J_2 17.5, $\frac{1}{2}COCH_2$), 4.34 (*c*, 1H, dd, J_1 2.5, J_2 7.7, αCH), 4.46 (*t*, 1H, dd, J_1 4.1, J_2 7.0, αCH) and 5.37 (*t* & *c*, s, 1H, NH); δ_C (74.76 MHz; C^2HCl_3) 21.39 (*c*, γCH_2), 23.87 (*t*, γCH_2), 27.49 (*t* & *c*, $C(\underline{C}H_3)_3$), 28.17 (*t*, βCH_2), 30.50 (*c*, βCH_2), 42.09 (*t* & *c*, $CO\underline{C}H_2$), 45.04 (*t*, δCH_2), 45.77 (*c*, δCH_2), 51.28 (*t*, OCH_3), 51.61 (*c*, OCH_3), 57.62 (*c*, αCH), 58.05 (*t*, αCH), 78.34 (*t* & *c*, $\underline{C}(CH_3)_3$), 155.00 (*t* & *c*, $CO_2{}^tBu$), 166.71 (*t*, $\underline{C}OCH_2$), 166.98 (*c*, $\underline{C}OCH_2$), 171.29 (*c*, $\underline{C}O_2CH_3$) and 171.57 (*t*, $\underline{C}O_2CH_3$); *m/z* (CI) 287 ($[M + H]^+$, 3%), 231 (18, $[M - C(CH_3)_3 + 2H]^+$), 213 (100, $[M - OC(CH_3)_3]^+$), 187 (23, $[M - CO_2C(CH_3)_3 + 2H]^+$) and 155 (48, $[M - CH_2NHCO_2{}^tBu + H]^+$).

N-tert-Butoxycarbonyl-glycyl-(2_S_)-prolinamide (23)

To a saturated solution of ammonia in dry methanol (50 cm^3) was added _N-tert_-butoxycarbonyl-glycyl-(2_S_)-proline methyl ester (**22**) (1.43 g, 5 mmol). The solution was placed in a tightly stoppered vessel and left at room temperature. When the reaction was complete as judged by TLC (10 days), the stoppered vessel was cooled to 0 °C, opened and nitrogen gas bubbled through at room temperature for 5 min. The solvent was then removed under reduced pressure to give the product as a white solid (1.34 g, 98%), m.p. 152-4 °C; (Found: C, 53.4; H, 8.05; N, 15.6. Calc. for $C_{12}H_{21}N_3O_4$: C, 53.15; H, 7.8; N, 15.5%); _m/z_ (Found: $[M + H]^+$, 272.1615. Calc. for $C_{12}H_{22}N_3O_4$: 272.1611); $[\alpha]_D^{22}$ -64.0 (_c_ 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3418 & 3381 (NH), 1688 (carbamate CO), 1673 (tertiary amide CO) and 1642 (primary amide CO); δ_H (200 MHz; C^2HCl_3) 1.40 (9H, s, $C(CH_3)_3$), 1.81-2.36 (4H, m, βCH_2 and γCH_2), 3.41 (1H, dd, J_1 8.8, J_2 17.5, $\frac{1}{2}\delta CH_2$), 3.57 (1H, dd, J_1 11.0, J_2 17.5, $\frac{1}{2}\delta CH_2$), 3.91 (2H, d, J 5.2, $COCH_2$), 4.52 (1H, m, αCH), 5.57 (1H, d, J 5.2, $CH_2N\underline{H}$), 5.94 (1H, s, $\frac{1}{2}CONH_2$) and 6.96 (1H, s, $\frac{1}{2}CONH_2$); δ_C (50.31 MHz; C^2HCl_3) 22.85 (_c_, γCH_2), 25.15 (_t_, γCH_2), 28.75 (_t_, βCH_2), 29.59 (_t_ & _c_, $C(\underline{C}H_3)_3$), 32.59 (_c_, βCH_2), 43.16 (_c_, $CO\underline{C}H_2$), 43.36 (_t_, $CO\underline{C}H_2$), 46.80 (_t_, δCH_2), 47.53 (_c_, δCH_2), 59.95 (_c_, αCH), 60.56 (_t_, αCH), 79.86 (_t_ & _c_, $\underline{C}(CH_3)_3$), 156.72 (_t_ & _c_, $CO_2{}^tBu$), 168.66 (_c_, $\underline{C}OCH_2$), 169.16 (_t_, $\underline{C}OCH_2$), 175.15 (_c_, $CONH_2$) and 175.39 (_t_, $CONH_2$); _m/z_ (CI) 272 ($[M + H]^+$, 32%), 216 (100, $[M - C(CH_3)_3 + 2H]^+$), 198 (26, $[M - OC(CH_3)_3]^+$), 172 (95, $[M - CO_2C(CH_3)_3 + 2H]^+$), 155 (31, $[M - NHCO_2C(CH_3)_3]^+$) and 70 (41, $[C_4H_8N]^+$).

Glycyl-(2_S_)-proline methyl ester trifluoroacetate (24)

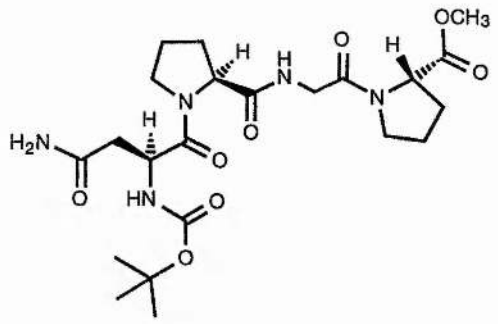
To a solution of trifluoroacetic acid (5.70 g, 50 mmol) in CH_2Cl_2 (25 cm^3) was added _N-tert_-butoxycarbonyl-glycyl-(2_S_)-proline methyl ester (**22**) (2.86 g, 10 mmol). The resultant solution was stirred at room temperature for 2 h and then the solvent and excess reagent removed under reduced pressure. The resulting oil was repeatedly redissolved in CH_2Cl_2 and the solvent removed under reduced pressure until no traces of TFA remained. The product was obtained as a clear oil (2.87 g, 100%), _m/z_ (Found: M^+, 187.1079. Calc. for $C_8H_{15}N_2O_3$: 187.1082); $[\alpha]_D^{22}$ -47.9 (_c_ 0.1 in MeOH); υ_{max} (thin film)/cm^{-1} 1783 (trifluoroacetate CO), 1742 (ester CO), 1666 (amide CO) and 706 (CF); δ_H (200 MHz; C^2HCl_3) 1.73-2.36 (_t_ & _c_, 4H, m, βCH_2 and γCH_2), 3.32-3.62 (_t_ & _c_, 2H, m, δCH_2), 3.68 (_t_, 3H, s, OCH_3), 3.74 (_c_, 3H, s, OCH_3), 3.94 (_t_ & _c_, 2H, s, $COCH_2$), 4.47 (_t_ & _c_, 1H, dd, J_1 3.4, J 8.2, αCH) and 7.84 (_t_ & _c_,

and 7.84 (*t* & *c*, 3H, s, H_3N^+); δ_C (50.31 MHz; C^2HCl_3) 22.41 (*c*, γCH_2), 24.72 (*t*, γCH_2), 29.22 (*t*, βCH_2), 31.10 (*c*, βCH_2), 41.14 (*t* & *c*, OCH_3), 46.58 (*t*, δCH_2), 47.31 (*c*, δCH_2), 52.87 (*t*, $CO\underline{C}H_2$), 53.28 (*c*, $CO\underline{C}H_2$), 59.26 (*c*, αCH), 59.71 (*t*, αCH), 116.39 (q, J_{CF} 1154, $CO\underline{C}F_3$), 161.65 (q, J_{CF} 147, $\underline{C}OCF_3$), 165.71 (*t*, $\underline{C}OCH_2$), 165.99 (*c*, $\underline{C}OCH_2$), 172.20 (*c*, $\underline{C}O_2CH_3$) and 172.84 (*t*, $\underline{C}O_2CH_3$); *m/z* (EI) 187 (M^+, 100%), 170 (7, $[M - NH_3]^+$), 156 (19, $[M - OCH_3]^+$), 130 (16, $[M - CO_2CH_3 + 2H]^+$) and 70 (8, $[C_4H_8N]^+$).

Glycyl-(2*S*)-prolinamide trifluoroacetate (25)

This compound was prepared in a manner identical with that for glycyl-(2*S*)-proline methyl ester trifluoroacetate (**24**), using *N*-*tert*-butoxycarbonyl-glycyl-(2*S*)-prolinamide (**23**) (1.36 cm^3, 5 mmol) to give the product as a clear oil (1.43 g, 100%), *m/z* (Found: M^+, 172.2069. Calc. for $C_7H_{14}N_3O_2$: 172.2065); $[\alpha]_D^{22}$ -41.3 (*c* 0.5 in MeOH); υ_{max} (thin film)/cm^{-1} 3050 (NH), 1781 (trifluoroacetate CO), 1667 (primary amide CO), 1656 (tertiary amide CO) and 711 (CF); δ_H (200 MHz; $C^2H_3O^2H$) 1.79-2.47 (4H, m, βCH_2 and γCH_2), 3.58 (2H, m, δCH_2), 3.89 (2H, s, $COCH_2$) and 4.41 (1H, dd, J_1 3.8, *J* 8.4, αCH); δ_C (50.31 MHz; $C^2H_3O^2H$) 23.60 (*c*, γCH_2), 25.66 (*t*, γCH_2), 31.08 (*t*, βCH_2), 33.41 (*c*, βCH_2), 41.62 (*c*, $COCH_2$), 41.85 (*t*, $COCH_2$), 47.85 (*t*, δCH_2), 48.63 (*c*, δCH_2), 61.34 (*c*, αCH), 62.13 (*t*, αCH), 117.93 (q, J_{CF} 1166, $CO\underline{C}F_3$), 162.44 (q, J_{CF} 143, $\underline{C}OCF_3$), 166.85 (*t* & *c*, $\underline{C}ONH_2$), 174.12 (*c*, $\underline{C}OCH_2$) and 174.87 (*t*, $\underline{C}OCH_2$); *m/z* (CI) 172 (M^+, 11%), 155 (100, $[M - NH_3]^+$), 111 (5, $[M - NH_3 - CONH_2]^+$), 98 (28, $[M - CH_2NH_3 - CONH_2 + H]^+$) and 70 (32, $[C_4H_8N]^+$).

N-tert-Butoxycarbonyl-(2_S_)-asparagyl-(2_S_)-prolyl-glycyl-(2_S_)-proline methyl ester (NP^1GP^2) (26)



To a solution of PyBOP (0.62 g, 1.2 mmol) in dry DMF (3 cm^3) was added _N_-methylmorpholine (0.13 cm^3, 1.2 mmol) and a solution of _N-tert_-butoxycarbonyl-(2_S_)-asparagyl-(2_S_)-proline (**19**) (0.33 g, 1 mmol) in dry THF/ dry DMF (10 cm^3, 1:1). A mixture of glycyl-(2_S_)-proline methyl ester trifluoroacetate (**24**) (0.30 g, 1 mmol) and _N_-methylmorpholine (0.12 cm^3, 1 mmol) in dry THF (5 cm^3) was then added and the solution stirred vigorously at room temperature under nitrogen for 50 min. The resultant white precipitate was filtered, washed with diethyl ether and dried to give the product as a white solid (0.38 g, 75%), m.p. 168-70 °C; _m/z_ (Found: $[M + H]^+$, 498.2559. Calc. for $C_{22}H_{36}N_5O_8$: 498.2563); $[\alpha]_D^{22}$ -131.0 (_c_ 0.1 in MeOH); υ_{max} (Nujol)/cm^{-1} 3355 & 3286 (NH), 1752 (ester CO), 1734 (carbamate CO), 1709 (tertiary amide CO), 1667 (tertiary amide CO), 1654 (secondary amide CO) and 1649 (primary amide CO); δ_H (200 MHz; C^2HCl_3) 1.42 (9H, s, $C(CH_3)_3$), 1.78-2.40 (8H, m, 2 x γCH_2 and 2 x βCH_2), 2.62 (1H, dd, J_1 15.8, J_2 3.2, $\frac{1}{2}C\underline{H}_2CONH_2$), 2.79 (1H, dd, J_1 15.8, J_2 10.0, $\frac{1}{2}C\underline{H}_2CONH_2$), 3.41-3.64 (2H, m, δCH_2^1), 3.69 (3H, s, OCH_3), 3.82 (2H, m, δCH_2^2), 4.03 (2H, d, _J_ 6.0, $C\underline{H}_2NH$), 4.45 (1H, dd, J_1 3.6, J_2 5.2, αCH^1), 4.64 (1H, d, _J_ 5.4, αCH^2), 4.83 (1H, m, $C\underline{H}CH_2CONH_2$), 5.59 (1H, s, $\frac{1}{2}CONH_2$), 5.92 (1H, d, _J_ 8.8, $N\underline{H}CHCH_2$), 6.78 (1H, s, $\frac{1}{2}CONH_2$) and 7.50 (1H, t, _J_ 6.0, $N\underline{H}CH_2$); δ_C (74.76 MHz; C^2HCl_3) 22.07 (_tc_, γCH_2^1 and $\gamma\underline{C}H_2^2$), 24.31 (_tt_, γCH_2^1), 24.54 (_tt_, γCH_2^2), 28.18 (_tt_ & _tc_, $C(CH_3)_3$), 28.82 (_tt_, βCH_2^1 and βCH_2^2), 31.10 (_tc_, βCH_2^1 and βCH_2^2), 37.88 (_tt_ & _tc_, $\underline{C}H_2CONH_2$), 41.53 (_tc_, $CONH\underline{C}H_2$), 41.64 (_tt_, $CONH\underline{C}H_2$), 46.04 (_tt_, δCH_2^1), 46.55 (_tc_, δCH_2^1), 47.30 (_tt_, δCH_2^2), 48.69 (_tc_, δCH_2^2), 52.25 (_tt_, αCH^1), 52.64 (_tc_, αCH^1), 58.56 (_tc_, $\alpha\underline{C}H^2$), 58.80 (_tt_, $\alpha\underline{C}H^2$), 60.15 (_tt_ & _tc_, $\underline{C}HCH_2CONH_2$), 79.63 (_tt_ & _tc_, $\underline{C}(CH_3)_3$), 155.16 (_tt_ & _tc_, $CO_2{}^tBu$), 167.40 (_tt_, $CONH_2$), 167.82 (_tc_, $CONH_2$), 171.52 (_tc_, $\underline{C}OCH_2NH$), 171.60 (_tt_, $\underline{C}OCH_2NH$), 171.78 (_tt_, $\underline{C}OCHNH$), 172.15 (_tc_, $\underline{C}OCHNH$), 172.52 (_tc_, $\underline{C}ONHCH_2$), 172.54 (_tt_, $\underline{C}ONHCH_2$) and 172.85 (_tt_ & _tc_, $\underline{C}O_2CH_3$); _m/z_ (CI) 498 ($[M + H]^+$, 41%), 398 (73,

[*M* - $CO_2C(CH_3)_3$ + 2H]$^+$), 380 (60, [*M* - $CO_2C(CH_3)_3$ - NH_2]$^+$), 284 (27, [Pro-Gly-Pro + H]$^+$), 212 (79, [Asn-Pro - OH]$^+$) and 187 (100, [Gly-Pro-OMe + H]$^+$).

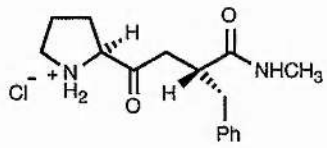
N-tert-Butoxycarbonyl-(2*S*)-asparagyl-(2*S*)-prolyl-glycyl-(2*S*)-prolinamide (NP^{1}GP^2) (27)

This compound was prepared in a manner identical with that for *N-tert*-butoxycarbonyl-(2*S*)-asparagyl-(2*S*)-prolyl-glycyl-(2*S*)-proline methyl ester (**26**), using glycyl-(2*S*)-prolinamide trifluoroacetate (**25**) (0.28 g, 1 mmol). The product was isolated as a white solid (0.79 g, 62%), m.p. 60-4 °C; *m/z* (Found: [*M* + H]$^+$, 483.2561. Calc. for $C_{21}H_{35}N_6O_7$: 483.2567); $[\alpha]_D^{22}$ -88.2 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3420 & 3250 (NH), 1736 (carbamate CO), 1705 (tertiary amide CO), 1661 (tertiary amide CO), 1649 (secondary amide CO), 1647 (primary amide CO) and 1638 (primary amide CO); δ_H (200 MHz; d_6-DMSO) 1.41 (9H, s, $C(CH_3)_3$), 1.81-2.32 (8H, m, 2 x γCH_2 and 2 x βCH_2), 2.53 (1H, dd, J_1 16.2, J_2 3.8 $\frac{1}{2}C\underline{H}_2CONH_2$), 2.79 (1H, dd, J_1 16.2, J_2 8.5, $\frac{1}{2}C\underline{H}_2CONH_2$), 3.43-3.76 (4H, m, 2 x δCH_2), 3.97 (1H, d, *J* 6.4, $\frac{1}{2}C\underline{H}_2NH$), 4.01 (1H, d, *J* 6.4, $\frac{1}{2}C\underline{H}_2NH$), 4.39 (1H, dd, J_1 3.4, J_2 6.1, αCH^1), 4.50 (1H, m, αCH^2), 4.72 (1H, m, $C\underline{H}CH_2CONH_2$), 6.34 (1H, s, $\frac{1}{2}CONH_2$), 6.46 (1H, d, *J* 8.4, $N\underline{H}CHCH_2$), 6.76 (1H, s, $\frac{1}{2}CONH_2$), 7.40 (1H, m, $N\underline{H}CH_2$) and 7.90 (1H, s, $\frac{1}{2}CONH_2$); δ_C (74.76 MHz; $C^2H_3O^2H$) 23.59 (*tc*, γCH_2^1 and $\gamma\underline{C}H_2^2$), 25.81 (*tt*, γCH_2^1), 26.99 (*tt*, γCH_2^2), 28.94 (*tt* & *tc*, $C(CH_3)_3$), 31.00 (*tc*, βCH_2^1), 31.99 (*tt*, βCH_2^2), 33.57 (*tc*, βCH_2^1 and βCH_2^2), 38.63 (*tt* & *tc*, $\underline{C}H_2CONH_2$), 42.88 (*tt*, $CONH\underline{C}H_2$), 43.10 (*tc*, $CONH\underline{C}H_2$), 47.87 (*tt*, δCH_2^1), 47.97 (*tc*, δCH_2^1), 48.17 (*tt*, δCH_2^2), 48.32 (*tc*, δCH_2^2), 52.86 (*tt* & *tc*, αCH^1), 61.24 (*tc*, $\alpha\underline{C}H^2$), 61.69 (*tt*, $\alpha\underline{C}H^2$), 61.79 (*tt*, $\underline{C}HCH_2CONH_2$), 61.97 (*tc*, $\underline{C}HCH_2CONH_2$), 80.73 (*tt* & *tc*, $\underline{C}(CH_3)_3$), 155.79 (*tt* & *tc*, $CO_2{}^tBu$), 169.96 (*tt*, $CH_2\underline{C}ONH_2$), 170.09 (*tc*, $CH_2\underline{C}ONH_2$), 173.29 (*tt*, $\underline{C}OCH_2NH$), 174.93 (*tc*, $\underline{C}OCH_2NH$), 175.44 (*tt*, $\underline{C}OCHNH$), 172.15 (*tc*, $\underline{C}OCHNH$), 172.52 (*tc*, $\underline{C}ONHCH_2$), 172.54 (*tt*, $\underline{C}ONHCH_2$), and 177.57 (*tt* & *tc*, $CH\underline{C}ONH_2$); *m/z* (CI) 483 ([*M* + H]$^+$, 23%), 483 (80, [*M* - $CO_2C(CH_3)_3$ + 2H]$^+$), 365 (42, [*M* - $CO_2C(CH_3)_3$ - NH_2]$^+$), 322 (19, [*M* - $CO_2C(CH_3)_3$ - NH_2 - $CONH_2$ + H]$^+$), 197 (64, [Asn-Pro - OH]$^+$) and 172 (100, [Gly-Pro-NH_2 + H]$^+$).

(2*S*)-Phenylalanine methylamide (31)

To a saturated solution of methylamine in dry methanol (200 cm^3) was added (2*S*)-phenyalanine methyl ester hydrochloride (2.16 g, 10 mmol). The mixture was placed in a tightly stoppered vessel and left at room temperature for 12 h. The solution was then cooled to 0 °C, opened carefully and nitrogen bubbled through for 10 min. The solvent was removed under reduced pressure to give a white solid. Water (100 cm^3) was added and the solution extracted with ethyl acetate (3 x 50 cm^3). The combined organic fractions were dried ($MgSO_4$) and the solvent removed under reduced pressure to give a white solid which was recrystallised from diethyl ether to give the product as colourless crystals (1.56 g, 88%), m.p. 160-2 °C; (Found: C, 67.75; H, 8.1; N, 16.0. Calc. for $C_{10}H_{14}N_2O$: C, 67.5 H, 7.9; N, 15.85%); *m/z* (Found: $[M + H]^+$, 179.1180. Calc. for $C_{10}H_{15}N_2O$: 179.1180); $[\alpha]_D^{22}$ +33.9 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3371 (amine NH), 3241 & 3291 (amide NH), 1645 (CO) and 747 & 701 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.42 (2H, s, NH_2), 2.79 (1H, dd, J_1 13.5, J_2 7.9, $\frac{1}{2}CH_2Ph$), 3.03 (1H, dd, J_1 13.5, J_2 5.1, $\frac{1}{2}CH_2Ph$), 3.64 (3H, s, CH_3), 3.68 (1H, dd, J_1 5.1, J_2 7.9, 2-CH) and 7.09-7.31 (5H, m, aromatic); δ_C (50.31 MHz; C^2HCl_3) 40.20 (CH_2Ph), 50.83 (CH_3), 54.91 (2-CH), 125.80 (*para* aromatic), 127.56 (*ortho* aromatic), 128.40 (*meta* aromatic), 136.55 (quat. aromatic) and 174.44 (CO); *m/z* (CI) 179 ($[M + H]^+$, 100%), 164 (3, $[M - CH_3 + H]^+$), 132 (5, $[M - NHCH_3 - NH_2]^+$) and 120 (17, $[M - CONHCH_3]^+$).

(2*S*)-Prolyl-(2*S*)-phenylalanine methylamide hydrochloride (32)



To a stirred solution of *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry THF (25 cm^3) was added *N-tert*-butoxycarbonyl-(2*S*)-proline (**30**) (2.15 g, 10 mmol) and the solution cooled to -15 °C. *iso*-Butyl chloroformate (1.36 cm^3, 10 mmol) was added and the resulting suspension stirred at -15 °C for 2 min. A solution of (2*S*)-phenylalanine methylamide (**31**) (1.78 g, 10 mmol) in dry THF (25 cm^3) was then added. The reaction mixture was allowed to warm to room temperature and left to stir for 1 h. The hydrochloride salts were filtered and the solvents removed under reduced pressure. The resultant clear oil was dissolved in CH_2Cl_2 (50 cm^3), washed with 0.5 mol dm^{-3} HCl (2 x 25 cm^3) and 5% sodium bicarbonate solution

(2 x 25 cm^3), dried ($MgSO_4$) and the solvent removed under reduced pressure to give a white solid. Hydrogen chloride gas was bubbled through a solution of the solid in ethylacetate (30 cm^3) for 20 min at 0 °C. The solvent was then removed under reduced pressure to give the product as a white solid which was filtered and recrystallized from methanol to give the product as colourless crystals (2.45 g, 78 %), m.p. 212-14 °C; *m/z* (Found: $[M + H]^+$, 276.1712. Calc. for $C_{15}H_{22}N_3O_2$: 276.1712); $[\alpha]_D^{22}$ -17.7 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3342 & 3318 (NH), 1680 (secondary amide CO) and 1643 (methylamide CO); δ_H (200 MHz; ^2H_2O) 1.86-2.02 (3H, m, $\frac{1}{2}\beta CH_2$ and γCH_2), 2.39 (1H, m, $\frac{1}{2}\beta CH_2$), 2.53 (3H, s, NCH_3), 2.96 (1H, dd, J_1 8.0, J_2 4.2, $\frac{1}{2}\beta CH_2Ph$), 3.10 (1H, dd, J_1 8.0, J_2 6.5, $\frac{1}{2}\beta CH_2Ph$), 3.23 (2H, m, δCH_2), 4.27 (1H, dd, J_1 4.2, J_2 6.5, αCH), 4.43 (1H, t, *J* 8.0, $C\underline{H}CH_2Ph$) and 7.21 (5H, m, aromatic); δ_C (50.3 MHz, ^2H_2O) 26.49 (γCH_2), 28.43 (NCH_3), 32.56 (βCH_2), 39.79 (CH_2Ph), 49.28 (δCH_2), 58.76 ($\underline{C}HCH_2Ph$), 62.26 (αCH), 130.06 (*para* aromatic), 131.92 (*ortho* aromatic), 131.55 (*meta* aromatic), 138.38 (quat. aromatic), 172.11 ($\underline{C}ONHCH$) and 175.77 ($\underline{C}ONCH_3$); *m/z* (CI) 276 ($[M + H]^+$, 100%), 245 (3, $[M - NH_2CH_3]^+$), 217 (2, $[M - CONH_2CH_3]^+$), 91 (2, $[PhCH_2]^+$) and 70 (33, $[C_4H_8N]^+$).

_N_-Benzyloxycarbonyl-glycyl-(2*S*)-prolyl-(2*S*)-phenylalanine methylamide (34)

To a solution of *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry THF (25 cm^3) was added (2*S*)-proline-(2*S*)-phenylalanine methylamide hydrochloride (**32**) (3.12 g, 10 mmol) and the resultant suspension stirred vigorously for 30 min at room temperature, (suspension A). Meanwhile, *N*-benzyloxycarbonyl-glycine (**33**) (2.09 g, 10 mmol) was added to a stirred solution of *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry THF (25 cm^3). The mixture was cooled to -15 °C and *iso*-butylchloroformate (1.36 cm^3, 10 mmol) was added. The resultant suspension was stirred at -15 °C for 2 min and then suspension A was added in one portion. The reaction mixture was allowed to warm to room temperature and then stirred for a futher 1 h. The hydrochloride salts were filtered off and the solvent removed under reduced pressure. The resultant clear oil was dissolved in CH_2Cl_2 (50 cm^3) and washed with 0.5 mol dm^{-3} HCl (2 x 25 cm^3) and 5% sodium bicarbonate solution (2 x 25 cm^3). The organic phase was dried ($MgSO_4$) and the solvent removed under reduced pressure to give the product as a white solid (3.88 g, 83%), m.p. 62-6 °C; *m/z* (Found: M^+, 467.2288. Calc. for $C_{25}H_{31}N_4O_5$: 467.2294); $[\alpha]_D^{22}$ -47.6 (*c* 0.5 in MeOH); υ_{max} (Nujol)/cm^{-1} 3303 (NH), 1696 (carbamate CO), 1661 (tertiary amide CO), 1652 (secondary amide CO) and 1648 (methylamide CO); δ_H (200 MHz; C^2HCl_3) 1.74-2.12 (4H, m, βCH_2 and γCH_2), 2.75 (3H, d, *J* 4.4, $NHC\underline{H}_3$), 2.98 (1H,

dd, J_1 9.5, J_2 14.3, $\frac{1}{2}$CHC$\underline{H}_2$Ph), 3.37-3.55 (3H, m, $\frac{1}{2}$CHC$\underline{H}_2$Ph and δCH_2), 3.73 (1H, dd, J_1 5.3, $J_2$16.6, $\frac{1}{2}$COCH_2), 3.88 (1H, dd, J_1 4.7, J_2 16.6, $\frac{1}{2}$COCH_2), 4.44 (1H, dd, J_1 3.3, J_2 7.5, αCH), 4.67 (1H, ddd, J_1 5.5, J_2 7.5, J_3 9.0, C$\underline{H}$CH_2Ph), 5.04 (1H, d, J 12.2, $\frac{1}{2}$OCH_2Ph), 5.16 (1H, d, J 12.2, $\frac{1}{2}$OC$\underline{H}_2$Ph), 5.62 (1H, t, J 5.0, N$\underline{H}$CH_2), 6.46 (1H, q, J 4.4, N$\underline{H}$CH_3), 6.92 (1H, d, J 9.5, N$\underline{H}$CHCH_2), 7.21 (5H, m, aromatic, CHCH_2Ph) and 7.36 (5H, s, aromatic, OCH_2Ph); δ_C (50.31 MHz; C^2HCl_3) 24.74 (γCH_2), 26.65 (NHCH_3), 29.09 (βCH_2), 37.57 (CH$\underline{C}H_2$Ph), 43.76 (CO$\underline{C}H_2$), 47.31 (δCH_2), 54.72 ($\underline{C}$HCH_2Ph), 61.42 (αCH), 67.54 ($CO_2\underline{C}H_2$Ph), 127.05 (*para* aromatic), 128.45 (*meta* aromatic), 128.67 (*para* aromatic), 128.76 (*ortho* aromatic), 128.98 (*meta* aromatic), 129.48 (*ortho* aromatic), 136.53 (quat. aromatic), 138.13 (quat. aromatic), 157.64 ($\underline{C}O_2CH_2$Ph), 170.01 ($\underline{C}OCH_2$), 171.70 ($\underline{C}$ONHCH_3) and 172.22 ($\underline{C}$ONHCHCH_2); *m/z* (CI) 467 ([*M* + H]$^+$, 22%), 376 (100, [*M* - PhCH_2 + H]$^+$), 359 (5, [*M* - PhCH_2O]$^+$), 277 (4, [*M* - CO_2CH_2Ph - $CONH_3$ + 3H]$^+$), 91 (3, [PhCH_2]$^+$) and 70 (9, [C_4H_8N]$^+$).

N-tert-Butoxycarbonyl-glycyl-(2*S*)-prolyl-(2*S*)-phenylananine methylamide (35)

This compound was prepared in a manner identical with that for *N*-benzyloxycarbonyl-glycyl-(2*S*)-prolyl-(2*S*)-phenylananine methylamide (**34**), using (2*S*)-proline-(2*S*)-phenylalanine methylamide hydrochloride (**32**) (1.12 cm^3, 10 mmol) and *N-tert*-butoxycarbonyl-glycine (**20**) (1.75 g, 10 mmol) to give the compound as a white solid (3.53 g, 82%), m.p. 75-7 °C; (Found: C, 59.9; H, 7.3; N, 12.7. Calc. for $C_{22}H_{32}N_4O_5$: C, 61.05; H, 7.45; N, 12.9%); *m/z* (Found: [*M* + H]$^+$, 433.2435. Calc. for $C_{22}H_{33}N_4O_5$: 433.2451); $[\alpha]_D^{22}$ -94.4 (*c* 0.5 in MeOH); υ_{max} (Nujol)/cm^{-1} 3312 (NH), 1675 (carbamate CO), 1662 (tertiary amide CO), 1658 (secondary amide CO) and 1641 (methylamide CO); δ_H (400 MHz; C^2HCl_3) 1.41 (*c*, 9H, s, $(CH_3)_3$), 1.43 (*t*, 9H, s, $(CH_3)_3$), 1.51-1.64 (*t* & *c*, 1H, m, $\frac{1}{2}$γCH_2), 1.77-1.93 (*t* & *c*, 2H, m, $\frac{1}{2}$βCH_2 and $\frac{1}{2}$γCH_2), 1.95-2.05 (*t* & *c*, 1H, m, $\frac{1}{2}$βCH_2), 2.64 (*c*, 3H, d, J 4.4, NHC$\underline{H}_3$), 2.73 (*t*, 3H, d, J 4.4, NHC$\underline{H}_3$), 2.93 (*t* & *c*, 1H, dd, J_1 9.6, J_2 14.2, $\frac{1}{2}$CH_2Ph), 3.27-3.37 (*t* & *c*, 2H, m, $\frac{1}{2}$CH_2Ph and $\frac{1}{2}$δCH_2), 3.42 (*t* & *c*, 1H, m, $\frac{1}{2}$δCH_2), 3.67 (*t* & *c*, 1H, dd, J_1 5.6, J_2 16.8, $\frac{1}{2}$COCH_2), 3.82 (*t* & *c*, 1H, dd, J_1 4.4, J_2 16.8, $\frac{1}{2}$COCH_2), 4.22 (*c*, 1H, dd, J_1 2.2, J_2 8.6, αCH), 4.42 (*t*, 1H, dd, J_1 2.4, J_2 8.3, αCH), 4.63 (*t* & *c*, 1H, ddd, J_1 4.7, J_2 8.5, J_3 8.5, C$\underline{H}$CH_2Ph), 5.37 (*t*, 1H, m, N$\underline{H}$CH_2), 5.45 (*c*, 1H, m, N$\underline{H}$CH_2), 6.54 (*t* & *c*, 1H, q, J 4.4, N$\underline{H}$CH_3), 6.92 (*c*, 1H, d, J 8.2, N$\underline{H}$CHCH_2Ph), 6.97 (*t*, 1H, d, J 8.2, N$\underline{H}$CHCH_2Ph) and 7.21 (*t* & *c*, 5H, m, aromatic); δ_C (74.76 MHz; C^2HCl_3) 22.52 (*c*, δCH_2), 24.15 (*t*, δCH_2), 26.17 (*t* & *c*, NHCH_3), 28.01 (*t*, βCH_2), 28.16 (*t*, C($\underline{C}H_3$)$_3$), 29.14 (*c*, βCH_2), 29.52 (*c*, C($\underline{C}H_3$)$_3$),

36.94 (*t*, CH_2Ph), 37.91 (*c*, CH_2Ph), 42.84 (*t* & *c*, $CO\underline{C}H_2$), 46.59 (*t*, δCH_2), 46.84 (*c*, δCH_2), 54.11 (*t*, $\underline{C}HCH_2Ph$), 54.97 (*c*, $\underline{C}HCH_2Ph$), 60.08 (*c*, αCH), 60.55 (*t*, αCH), 80.04 (*c*, $\underline{C}(CH_3)_3$), 80.10 (*t*, $\underline{C}(CH_3)_3$), 126.58 (*t* & *c*, *para* aromatic), 128.25 (*t* & *c*, *meta* aromatic), 128.86 (*t* & *c*, *ortho* aromatic), 136.66 (*c*, quat. aromatic), 137.42 (*t*, quat. aromatic), 156.06 (*t*, $CO_2{}^tBu$), 156.40 (*c*, $CO_2{}^tBu$), 168.54 (*c*, $\underline{C}ONHCH_3$), 169.29 (*t*, $\underline{C}ONHCH_3$), 170.67 (*t* & *c*, $\underline{C}ONHCHCH_2$) and 171.19 (*t* & *c*, $\underline{C}OCH_2$); *m/z* (FAB) 433 ($[M + H]^+$, 100%), 401 (4, $[M + H - NHCH_3]^+$), 355 (15, $[M - Ph]^+$), 333 (96, $[M + 2H - CO_2{}^tBu]^+$), 302 (7, $[M - CH_2NHCO_2{}^tBu]^+$), 274 (15, $[M - COCH_2NHCO_2{}^tBu]^+$) and 255 (8, $[M - NHCHCH_2PhCONHCH_3]^+$).

_N_-Bromoacetyl-(2_S_)-proline methyl ester (38)

To a solution of *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry THF (25 cm^3) was added bromoacetic acid (**37**) (1.39 g, 10 mmol) and the solution cooled to -15 °C. *iso*-Butyl chloroformate (1.36 cm^3, 10 mmol) added with stirring and the resulting suspension stirred at -15 °C for 2 min. A mixture of (2*S*)-proline methyl ester hydrochloride (**21**) (1.66 g, 10 mmol) and *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry DMF (5 cm^3) was then added in one portion to the cold suspension. The reaction mixture was allowed to warm to room temperature and left stirring overnight. The hydrochloride salts were filtered off and the solvents removed under reduced pressure. The resulting brown oil was dissolved in CH_2Cl_2, washed with 0.5 mol dm^{-3} HCl (2 x 20 cm^3) and 5% sodium bicarbonate solution (2 x 20 cm^3), dried ($MgSO_4$) and the solvent removed under reduced pressure to yield the product as a pale yellow oil. Purification by silica column chromatography using ethyl acetate/ light petroleum (1:1) as the eluent gave the pure product as a clear oil (2.07 g, 83%), *m/z* (Found: $[M + H]^+$, 250.0080. Calc. for $C_8H_{13}{}^{79}BrNO_3$: 250.0079); $[\alpha]_D^{22}$ -96.7 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 1741 (ester CO), 1655 (amide CO) and 1448 (C-O); δ_H (200 MHz; C^2HCl_3) 1.81-2.28 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.48-3.73 (*t* & *c*, 2H, m, δCH_2), 3.57 (*t*, 3H, s, OCH_3), 3.62 (*c*, 3H, s, OCH_3), 3.73 (*c*, 2H, d, *J* 1.4, $COCH_2$), 3.96 (*t*, 2H, d, *J* 1.4, $COCH_2$), 4.33 (*t*, 1H, dd, J_1 4.0, J_2 8.2, αCH) and 4.45 (*c*, 1H, dd, J_1 3.4, J_2 7.2, αCH); δ_C (50.31 MHz; C^2HCl_3) 22.63 (*c*, γCH_2), 25.24 (*t*, γCH_2), 29.47 (*t*, βCH_2), 31.55 (*c*, βCH_2), 42.38 (*t* & *c*, CH_2Br), 47.44 (*t* & *c*, δCH_2), 52.65 (*t*, αCH), 53.32 (*c*, αCH), 59.61 (*t* & *c*, OCH_3), 165.49 (*t*, NCO), 165.59 (*c*, NCO) and 172.56 (*t* & *c*, $\underline{C}O_2CH_3$); *m/z* (EI) 252 & 250 (M^+, 3%), 170 (40, $[M - Br]^+$), 156 (2, $[M - Br - CH_3 + H]^+$), 139 (14, $[M - Br - OCH_3]^+$) and 70 (100, $[C_4H_8N]^+$).

(9a*S*)-2,3,7,8,9,9a-Hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (39)

To a solution of hydrazine hydrate (1.5 g, 30 mmol) in ethanol (50 cm^3) was added *N*-bromoacetyl-(2*S*)-proline methyl ester (**38**) (2.50 g, 10 mmol). The resultant solution was refluxed for 2 h and allowed to cool to room temperature whereupon colourless crystals of the product were formed (1.43 g, 57%), m.p. 265-6 °C; *m/z* (Found: $[M + H]^+$, 170.0930. Calc. for $C_7H_{12}N_3O_2$: 170.0930); $[\alpha]_D^{22}$ -74.4 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3403 (NH), 1664 (secondary amide CO) and 1637 (tertiary amide CO); δ_H (200 MHz; ^2H_2O) 1.83-2.25 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.41 (1H, m, $\frac{1}{2}\beta CH_2$), 3.42-3.71 (2H, m, δCH_2), 4.11 (1H, d, *J* 16.5, $\frac{1}{2}COCH_2$), 4.45-4.59 (1H, m, αCH) and 4.59 (1H, d, *J* 16.5, $\frac{1}{2}COCH_2$); δ_H (200 MHz; d_6-DMSO) 1.77-2.10 (*A* & *B*, 3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.14-2.41 (*A* & *B*, 1H, m, $\frac{1}{2}\beta CH_2$), 3.34-3.62 (*A* & *B*, 2H, m, δCH_2), 3.89 (*A*, 1H, d, *J* 15.1, $\frac{1}{2}COCH_2$), 4.11 (*B*, 1H, d, *J* 16.0, $\frac{1}{2}COCH_2$), 4.40 (*A*, 1H, d, *J* 15.1, $\frac{1}{2}COCH_2$), 4.42 (*A* & *B*, 1H, m, αCH) and 4.61 (*B*, 1H, d, *J* 16.0, $\frac{1}{2}COCH_2$); δ_C (50.31 MHz; ^2H_2O) 24.71 (*A*, γCH_2), 24.85 (*B*, γCH_2), 30.56 (*B*, βCH_2), 30.72 (*A*, βCH_2), 48.39 (*A*, $CO\underline{C}H_2$), 48.47 (*B*, $CO\underline{C}H_2$), 54.51 (*A*, δCH_2), 54.90 (*B*, δCH_2), 61.44 (*A*, αCH), 61.59 (*B*, αCH), 166.61 (*A*, $\underline{C}OCH_2$), 166.80 (*B*, $\underline{C}OCH_2$), 170.47 (*B*, CONH), 170.96 (*A*, CONH); δ_C (50.31 MHz; d_6-DMSO) 25.99 (*A*, γCH_2), 26.11 (*B*, γCH_2), 31.89 (*B*, βCH_2), 32.12 (*A*, βCH_2), 48.75 (*A* & *B*, $CO\underline{C}H_2$), 56.32 (*B*, δCH_2), 56.53 (*A*, δCH_2), 61.78 (*A*, αCH), 61.88 (*B*, αCH), 165.91 (*A*, $\underline{C}OCH_2$), 166.17 (*B*, $\underline{C}OCH_2$), 169.66 (*B*, CONH) and 170.44 (*A*, CONH); *m/z* (CI) 170 ($[M + H]^+$, 100%), 155 (12, $[M - NH]^+$), 137 (7, $[M - 2NH_3]^+$) and 70 (5, $[C_4H_8N]^+$).

(2*S*)-Bromopropionic acid (44)

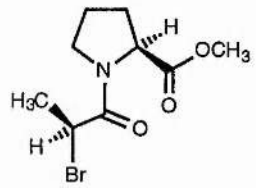
(2*S*)-Alanine (**42**) (4.0 g, 45 mmol) was added to a saturated solution of potassium bromide (10 cm^3). Hydrogen bromide (15 cm^3 of a 48% solution) as added dropwise and then the reaction mixture cooled to 0 °C and sodium nitrite (6.21 g, 90 mmol) added over 1 h. The reaction mixture was maintained below 5 °C for a further hour and then allowed to warm to room temperature. The resulting solution was extracted with diethyl ether (3 x 25 cm^3) and the combined ethereal extracts were then dried

($MgSO_4$) and then concentrated under reduced pressure to give a pale yellow oil which was distilled to give pure (2*S*)-bromopropionic acid (5.87 g, 85%), b.p. 100 °C/ 10 mmHg (lit.,[207] 87-8 °C/ 0.5 mmHg); *m/z* (Found: [*M* + H]$^+$, 151.9478. $C_3H_5{}^{79}BrO_2$ requires 151.9478); $[\alpha]_D^{22}$ -31.8 (*c* 1.0 in MeOH) {lit.,[207] -34.0 (*c* 1.0 in MeOH)}; υ_{max} (thin film)/cm^{-1} 1724 (CO) and 1245 (C-O); δ_H (200 MHz; C^2HCl_3) 1.88 (3H, d, *J* 6.7, CH_3), 4.41 (1H, d, *J* 6.7, 2-CH) and 10.75 (1H, s, CO_2H); δ_C (50.31 MHz; C^2HCl_3) 21.02 (CH_3), 39.22 (2-CH) and 176.26 (CO_2H); *m/z* (EI) 154 & 152 (*M*$^+$, 5%), 73 (21, [*M* - Br]$^+$), 45 (32, [*M* - COBr]$^+$) and 27 (100, [*M* - $HBrCO_2H$]$^+$).

(2*R*)-Bromopropionic acid (45)

This compound was prepared in a manner identical with that for the (2*S*) epimer (**44**), using (2*R*)-alanine (**43**) to give a pale yellow oil which was distilled under reduced pressure to give the pure product as a clear oil (5.73 g, 83%); b.p. 100 °C/ 10 mmHg; $[\alpha]_D^{22}$ +29.9 (*c* 1.0 in MeOH); All other spectral data were identical to that quoted above.

***N*-(2*S*)-Bromopropanoyl-(2*S*)-proline methyl ester (47)**



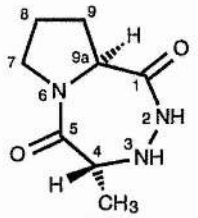
To a stirred solution of (2*S*)-bromopropionic acid (**44**) (1.53 g, 10 mmol) in dry THF (20 cm^3) at -40 °C was added *N*-methylmorpholine (1.12 cm^3, 10 mmol). *iso*-Butyl chloroformate (1.36 cm^3, 10 mmol) was immediately added and the resultant suspension stirred at -40 °C for 2 min. A mixture of (2*S*)-proline methyl ester hydrochloride (**21**) (1.66 g, 10 mmol) and *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry DMF (5 cm^3) was then added. The reaction mixture was allowed to warm up to room temperature and left to stir for a futher 1 h. The hydrochloride salts were filtered off and the solvents removed under reduced pressure. The resultant clear oil was dissolved in CH_2Cl_2 (25 cm^3) and washed with 0.5 mol dm^{-3} HCl (2 x 15 cm^3) and 5% sodium carbonate solution (2 x 15 cm^3). The organic phase was dried ($MgSO_4$) and the solvent removed under reduced pressure to yield a white solid.

Recrystallisation from diethyl ether/ light petroleum gave the product as colourless crystals (2.19 g, 83%), m.p. 116-8 °C; (Found: C, 41.85; H, 5.4; N, 5.35. Calc. for $C_9H_{14}BrNO_3$: C, 42.0; H, 5.35; N, 5.3%); *m/z* (Found: $[M + H]^+$, 264.0235. Calc. for $C_9H_{15}{}^{79}BrNO_3$: 264.0235); $[\alpha]_D^{22}$ +127.3 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 1752 (ester CO), 1648 (amide CO) and 1197 (C-O); δ_H (400 MHz; C^2HCl_3) 1.76 (*t*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.77 (*c*, 3H, d, *J* 6.7, $CHC\underline{H}_3$), 1.80-2.28 (*t* & *c*, 4H, m, βCH_2 & γCH_2), 3.45-3.64 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 3.69 (*t*, 3H, s, OCH_3), 3.73 (*c*, 3H, s, OCH_3), 3.79-3.86 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 4.18 (*t*, 1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.40 (*c*, 1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.45 (*t*, 1H, dd, J_1 4.1, J_2 8.6, αCH) and 4.61 (*c*, 1H, dd, J_1 4.1, J_2 8.6, αCH); δ_C (100 MHz; C^2HCl_3) 20.82 (*t*, $CH\underline{C}H_3$), 20.97 (*c*, $CH\underline{C}H_3$), 21.53 (*c*, γCH_2), 24.56 (*t*, γCH_2), 28.82 (*t*, βCH_2), 30.77 (*c*, βCH_2), 39.25 (*t*, $\underline{C}HCH_3$), 39.97 (*c*, $\underline{C}HCH_3$), 46.69 (*c*, δCH_2), 46.81 (*t*, δCH_2), 52.03 (*t*, OCH_3), 52.65 (*c*, OCH_3), 58.99 (*t*, αCH), 59.05 (*c*, αCH), 167.56 (*t*, $\underline{C}OCHCH_3$), 167.99 (*c*, $\underline{C}OCHCH_3$), 171.87 (*t*, $\underline{C}O_2CH_3$) and 172.05 (*c*, $\underline{C}O_2CH_3$); *m/z* (FAB) 264 & 266 (M^+, 100%), 204 & 206 (38, $[M - HCOOCH_3]^+$), 184 (35, $[M - Br]^+$) and 128 (62, $[M - COCHCH_3Br]^+$).

_N_-(2_R_)-Bromopropanoyl-(2_S_)-proline methyl ester (48)

This compound was prepared in a manner identical with that for the (2*S*, 2*S*) methyl ester (**47**), using (2*R*)-bromopropionic acid (**45**) (1.53 g, 10 mmol) to give the product as colourless crystals (2.09 g, 79%), m.p. 113-5 °C; (Found: C, 41.95; H, 5.35; N, 5.3. Calc. for $C_9H_{14}BrNO_3$: C, 42.0; H, 5.35; N, 5.3%); *m/z* (Found: $[M + H]^+$, 264.0232. Calc. for $C_9H_{15}{}^{79}BrNO_3$: 264.0235); $[\alpha]_D^{22}$ +113.6 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 1752 (ester CO), 1646 (amide CO) and 1197 (C-O); δ_H (400 MHz; C^2HCl_3) 1.62 (*c*, 3H, d, *J* 6.4, $CHC\underline{H}_3$), 1.64 (*t*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.82-2.31 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.44-3.66 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 3.71 (*t*, 3H, s, OCH_3), 3.75 (*c*, 3H, s, OCH_3), 3.82-3.94 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 4.22 (*c*, 1H, q, *J* 6.5, $C\underline{H}CH_3$), 4.44 (*t*, 1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.47 (*t*, 1H, dd, J_1 4.3, J_2 8.6, αCH) and 4.69 (*c*, 1H, dd, J_1 4.0, J_2 6.7, αCH); δ_C (100 MHz; C^2HCl_3) 20.21 (*t*, $CH\underline{C}H_3$), 20.35 (*c*, $CH\underline{C}H_3$), 22.21 (*c*, γCH_2), 24.67 (*t*, γCH_2), 28.82 (*t*, βCH_2), 30.89 (*c*, βCH_2), 39.27 (*t*, $\underline{C}HCH_3$), 40.02 (*c*, $\underline{C}HCH_3$), 46.70 (*c*, δCH_2), 46.75 (*t*, δCH_2), 50.69 (*t*, OCH_3), 50.81 (*c*, OCH_3), 58.87 (*c*, αCH), 59.04 (*t*, αCH), 167.44 (*t*, $\underline{C}OCHCH_3$), 167.79 (*c*, $\underline{C}OCHCH_3$), 171.92 (*t*, $\underline{C}O_2CH_3$) and 172.24 (*c*, $\underline{C}O_2CH_3$); *m/z* (FAB) 264 & 266 (M^+, 95%), 204 & 206 (62, $[M - HCOOCH_3]^+$), 184 (41, $[M - Br]^+$) and 128 (100, $[M - COCHCH_3Br]^+$).

(4*R*, 9a*S*)-4-Methyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (49)



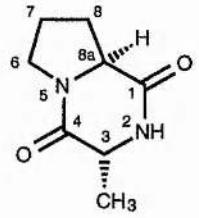
To a solution of hydrazine hydrate (0.60 g, 12 mmol) in ethanol (20 cm^3) was added *N*-(2*S*)-bromopropionyl-(2*S*)-proline methyl ester (**47**) (2.64 g, 10 mmol). The resultant solution was refluxed for 16 h, allowed to cool slowly and the hydrazine hydrobromide precipitate removed by filtration. The solvents were removed from the filtrate under reduced pressure to give the product as a colourless oil (1.81 g, 98%); *m/z* (Found: $[M + H]^+$, 184.1083. Calc. for $C_8H_{14}N_3O_2$: 184.1086); $[\alpha]_D^{22}$ +39.0 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 1665 (secondary amide CO) and 1639 (tertiary amide CO); δ_H (200 MHz; $C^2H_3O^2H$) 1.53 (3H, d, *J* 7.0, $CHC\underline{H}_3$), 1.86-2.15 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.39 (1H, m, $\frac{1}{2}\beta CH_2$), 3.44-3.72 (2H, m, δCH_2), 4.12 (1H, q, *J* 7.0, $C\underline{H}CH_3$) and 4.38 (1H, dd, J_1 7.5, J_2 7.5, αCH); δ_C (50.31 MHz; $C^2H_3O^2H$) 17.01 ($CH\underline{C}H_3$), 23.65 (γCH_2), 30.57 (βCH_2), 46.97 (δCH_2), 59.60 (αCH), 63.69 ($\underline{C}HCH_3$), 167.85 ($\underline{C}OCHCH_3$) and 168.49 (CONH); *m/z* (CI) 184 ($[M + H]^+$, 100%), 169 (20, $[M - CH_3 + H]^+$), 154 (8, $[M - NHCH_3 + H]^+$), 138 (4, $[M - NHNH - CH_3]^+$), 126 (5, $[M - NHNHCHCH_3 + H]^+$) and 70 (10, $[C_4H_8N]^+$).

(4*S*, 9a*S*)-4-Methyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (50)

This compound was prepared in a manner identical with that for the (4*R*, 9a*S*) diastereomer (**49**), using *N*-(2*R*)-bromo-propionyl-(2*S*)-proline methyl ester (**48**) (2.64 g, 10 mmol) to give the product as a colourless oil (1.75 g, 96%); *m/z* (Found: $[M + H]^+$, 184.1082. Calc. for $C_8H_{14}N_3O_2$: 184.1086); $[\alpha]_D^{22}$ +28.6 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 1658 (secondary amide CO) and 1645 (tertiary amide CO); δ_H (200 MHz; $C^2H_3O^2H$) 1.72 (3H, d, *J* 76.8 $CHC\underline{H}_3$), 1.93-2.29 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.52 (1H, m, $\frac{1}{2}\beta CH$), 3.54-3.86 (2H, m, δCH_2), 4.35 (1H, q, *J* 6.8, $C\underline{H}CH_3$) and 4.38 (1H, m, αCH); δ_C (50.31 MHz; $C^2H_3O^2H$) 18.36 ($CH\underline{C}H_3$), 24.72 (γCH_2), 31.43 (βCH_2), 48.55 (δCH_2), 61.24 (αCH), 64.54 ($\underline{C}HCH_3$), 167.49 ($\underline{C}OCHCH_3$) and 170.68 (CONH); *m/z* (CI) 184 ($[M + H]^+$, 100%), 169 (31, $[M - CH_3 + H]^+$), 154 (9,

[*M* - $NHNHCH_3$ + H]$^+$), 130 (4, [*M* - $COCHCH_3$ + 3H]$^+$), 126 (6, [*M* - $NHNHCHCH_3$ + H]$^+$) and 70 (15, $[C_4H_8N]^+$).

(3*R*, 8a*S*)-3-Methyl-2,3,6,7,8,8a-hexahydropyrrolo-[1,2-a]-pyrazine-1,4-dione (53)



To a solution of *N*-benzyloxycarbonyl-(2*R*)-alanyl-(2*S*)-prolinamide (**62**) (0.64 g, 2 mmol) in methanol (30 cm^3) was added 10% palladium on activated charcoal (30 mg) and the vessel flushed with hydrogen gas. The resultant suspension was stirred vigorously under an atmosphere of hydrogen for 12 h. The mixture was then filtered through a prewashed celite pad and the solvent removed under reduced pressure to give a white solid which was recrystallised from methanol to give the product as colourless crystals (0.32 g, 93%), m.p. 128-30 °C (lit.,[203] 127-9 °C); (Found: C, 57.4; H, 7.35; N, 16.65; *M*$^+$, 168.1902. $C_8H_{12}N_2O_2$ requires C, 57.15 H, 7.2 N, 16.65%; *M*$^+$, 168.1899); $[\alpha]_D^{22}$ -175.2 (*c* 1.0 in EtOH) {lit.,[203] -182.3 (*c* 1.0 in EtOH)}; υ_{max} (Nujol)/cm^{-1} 3213 & 3162 (NH), 1687 (tertiary amide CO) and 1635 (secondary amide CO); δ_H (200 MHz; $C^2H_3O^2H$) 1.41 (3H, d, *J* 7.0, CH_3), 1.84-2.13 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.63-2.84 (1H, m, $\frac{1}{2}\beta CH_2$), 3.42-3.70 (2H, m, δCH_2), 3.92 (1H, q, *J* 7.1, $C\underline{H}CH_3$) and 4.28 (1H, dd, J_1 6.5, J_2 9.7, αCH); δ_C (74.76 MHz; $C^2H_3O^2H$) 20.19 (CH_3), 23.23 (γCH_2), 30.21 (βCH_2), 46.87 (δCH_2), 54.71 ($\underline{C}HCH_3$), 59.46 (αCH), 169.34 ($\underline{C}OCHCH_3$) and 171.39 (CONH); *m/z* (EI) 168 (*M*$^+$, 71%), 140 (8, [*M* - $CHCH_3$]$^+$), 125 (28, [*M* - $NHCHCH_3$]$^+$), 112 (8, [*M* - $COCHCH_3$]$^+$), 97 (41, [*M* - $CONHCHCH_3$]$^+$) and 70 (100, $[C_4H_8N]^+$).

(3*S*, 8a*S*)-3-Methyl-2,3,6,7,8,8a-hexahydropyrrolo-[1,2-a]-pyrazine-1,4-dione (54)

This compound was prepared in a manner identical with that for the hexahydropyrrolo-[1,2-a]-pyrazine-1,4-dione (**53**), using *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (**61**) (0.638 g, 2 mmol) to give a white solid which was recrystallised from methanol to give the product as colourless crystals (0.31 g, 91%),

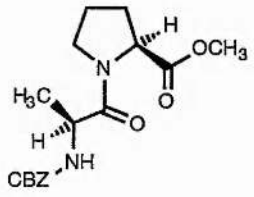
m.p. 152-6 °C (lit.,[203] 153-6 °C); (Found: C, 57.3; H, 7.25; N, 16.7; M^+, 168.1900. Calc. for $C_8H_{12}N_2O_2$: C, 57.15; H, 7.2; N, 16.65%; M^+, 168.1899); $[\alpha]_D^{22}$ -158.4 (*c* 1.0 in EtOH) {lit.,[203] -160.0 (*c* 1.0 in EtOH)}; υ_{max} (Nujol)/cm^{-1} 3287 (NH), 1687 (tertiary amide CO) and 1654 (secondary amide CO); δ_H (200 MHz; C^2HCl_3) 1.41 (3H, d, *J* 6.6, CH_3), 1.68-2.39 (4H, m, βCH_2 and γCH_2), 3.50 (2H, m, δCH_2), 4.08 (2H, m, αCH and C<u>H</u>CH_3) and 7.42 (1H, s, NH); δ_C (74.76 MHz; $C^2H_3O^2H$) 16.05 (CH_3), 23.89 (γCH_2), 29.43 (βCH_2), 46.48 (δCH_2), 52.33 (<u>C</u>HCH_3), 60.70 (αCH), 169.21 (<u>C</u>O$CHCH_3$) and 172.78 (CONH); *m/z* (EI) 168 (M^+, 43%), 140 (18, [*M* - $NHCH_3$ + 2H]$^+$), 125 (35, [*M* - $NHCHCH_3$]$^+$), 97 (34, [*M* - $CONHCHCH_3$]$^+$) and 70 (100, $[C_4H_8N]^+$).

_N_-Benzyloxycarbonyl-glycyl-(2*S*)-proline methyl ester (57)

To a solution of *N*-benzyloxycarbonyl-glycine (**33**) (2.09 g, 10 mmol) in dry THF (20 cm^3) was added *N*-methylmorpholine (1.12 cm^3, 10 mmol) and the solution cooled to -15 °C. *iso*-Butyl chloroformate (1.36 cm^3, 10 mmol) was added with stirring and the solution left stirred at -15 °C for 2 min. A solution of (2*S*)-proline methyl ester hydrochloride (**21**) (1.66 g, 10 mmol) and *N*-methylmorpholine (1.12 cm^3, 10 mmol) in dry DMF (5 cm^3) was then added. The reaction mixture was allowed to warm up to room temperature and left to stir for 2 h. The hydrochloride salts were filtered off and the solvents removed under reduced pressure. The resultant oily residue was dissolved in CH_2Cl_2 (25 cm^3) and washed with 0.5 mol dm^{-3} HCl (2 x 15 cm^3) and 5% sodium carbonate solution (2 x 15 cm^3). The organic phase was then dried ($MgSO_4$) and the solvent removed under reduced pressure to give a pale yellow oil. The crude material was purified by silica column chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a clear oil (2.42 g, 76%), *m/z* (Found: [*M* + H]$^+$, 321.1453. Calc. for $C_{16}H_{21}N_2O_5$: 321.1450); $[\alpha]_D^{22}$ -68.1 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3336 (NH), 1732 (ester CO), 1711 (carbamate CO), 1652 (amide CO) and 736 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.81-2.33 (*t* & *c*, 4H, m, βCH_2 and γCH_2) 3.35-3.69 (*t* & *c*, 2H, m, δCH_2), 3.73 (*t*, 3H, s, OCH_3), 3.76 (*c*, 3H, s, OCH_3), 4.02 (*t* & *c*, 2H, d, *J* 2.6, $COCH_2$), 4.40 (*t*, 1H, dd, J_1 4.0, J_2 6.6, αCH), 4.53 (*c*, 1H, dd, J_1 3.8, J_2 7.0, αCH), 5.12 (*t* & *c*, 2H, s, CH_2Ph), 5.74 (*t* & *c*, 1H, t, *J* 2.6, N<u>H</u>CH_2) and 7.35 (*t* & *c*, 5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 22.52 (*c*, γCH_2), 24.97 (*t*, γCH_2), 29.33 (*t*, βCH_2), 31.66 (*c*, βCH_2), 43.52 (*c*, OCH_3), 43.66 (*t*, OCH_3), 46.20 (*t*, δCH_2), 47.00 (*c*, δCH_2), 52.68 (*t*, $COCH_2$), 53.15 (*c*, $COCH_2$), 58.80 (*c*, αCH), 59.20 (*t*, αCH), 128.28 (*t* & *c*, *para* aromatic), 128.37 (*t* & *c*, *ortho* aromatic), 128.81 (*t* & *c*, *meta* aromatic), 136.87

(*t* & *c*, quat. aromatic), 156.63 (*t* & *c*, $\underline{C}O_2CH_2Ph$), 167.41 (*t*, $\underline{C}O_2CH_2$), 167.69 (*c*, $\underline{C}O_2CH_2$), 172.29 (*c*, $\underline{C}O_2CH_3$) and 172.68 (*t*, $\underline{C}O_2CH_3$); *m/z* (CI) 321 ([*M* + H]$^+$, 100%), 230 (10, [*M* - CH_2Ph + H]$^+$), 213 (28, [*M* - OCH_2Ph]$^+$), 187 (36, [*M* - CO_2CH_2Ph + 2H]$^+$), 157 (63, [*M* - CO_2CH_2Ph - OCH_3 + 3H]$^+$) and 91 (41, [CH_2Ph]$^+$).

_N_-Benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (58)



This compound was prepared in a manner identical with that for the *N*-benzyloxycarbonyl-glycyl-(2*S*)-proline methyl ester (**57**), using *N*-benzyloxycarbonyl-(2*S*)-alanine (**55**) (2.23 g, 10 mmol) instead of *N*-benzyloxycarbonyl-glycine (**33**) to give a yellow oil which was purified by silica chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a clear oil (2.89 g, 84%), *m/z* (Found: [*M* + H]$^+$, 335.1605. Calc. for $C_{17}H_{23}N_2O_5$: 335.1606); $[\alpha]_D^{22}$ -86.5 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 2982 & 2953 (NH), 1746 (ester CO), 1716 (carbamate CO), 1654 (amide CO) and 669 & 666 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.11 (*t* & *c*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.49-2.02 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.21-3.45 (*t* & *c*, 2H, m, δCH_2), 3.40 (*t* & *c*, 3H, s, OCH_3), 4.24 (*t* & *c*, 1H, m, αCH), 4.26 (*t* & *c*, 1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.82 (*t* & *c*, 2H, s, CH_2Ph), 5.97 (*c*, 1H, d, *J* 8.6, NH), 6.13 (*t*, 1H, d, *J* 7.8, NH) and 7.06 (*t* & *c*, 5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 18.03 (*t*, $CH\underline{C}H_3$), 19.06 (*c*, $CH\underline{C}H_3$), 22.91 (*c*, γCH_2), 25.14 (*t*, γCH_2), 29.09 (*t*, βCH_2), 31.47 (*c*, βCH_2), 46.75 (*c*, δCH_2), 46.96 (*t*, δCH_2), 48.56 (*t*, $\underline{C}HCH_3$), 48.77 (*c*, $\underline{C}HCH_3$), 52.30 (*t*, OCH_3), 53.87 (*c*, OCH_3), 58.97 (*t*, αCH), 59.29 (*c*, αCH), 66.67 (*t*, CH_2Ph), 67.05 (*c*, CH_2Ph), 128.14 (*t* & *c*, *ortho* aromatic), 128.68 (*t* & *c*, *para* aromatic), 129.36 (*t* & *c*, *meta* aromatic), 137.05 (*t* & *c*, quat. aromatic), 155.73 (*c*, $\underline{C}O_2CH_2Ph$), 156.13 (*t*, $\underline{C}O_2CH_2Ph$), 171.65 (*t* & *c*, $\underline{C}OCHCH_3$), 172.51 (*c*, $\underline{C}O_2CH_3$) and 172.62 (*t*, $\underline{C}O_2CH_3$); *m/z* (CI) 335 ([*M* + H]$^+$, 100%), 319 (7, [*M* - CH_3]$^+$), 291 (13, [*M* - $COCH_3$]$^+$), 227 (12, [*M* - OCH_2Ph]$^+$) and 91 (8, [$PhCH_2$]$^+$).

_N_-Benzyloxycarbonyl-(2_R_)-alanyl-(2_S_)-proline methyl ester (59)

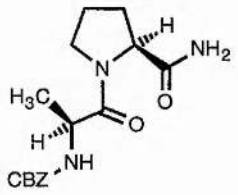
This compound was prepared in a manner identical with that for the methyl ester (**58**), using _N_-benzyloxycarbonyl-(2_R_)-alanine (**56**) (2.23 g, 10 mmol) instead of _N_ benzyloxycarbonyl-(2_S_)-alanine (**55**) to give a yellow oil which was purified by silica column chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a clear oil (2.40 g, 72%), _m/z_ (Found: $[M + H]^+$, 335.1614. Calc. for $C_{17}H_{23}N_2O_5$: 335.1607); $[\alpha]_D^{22}$ -22.9 (_c_ 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 2982 & 2955 (NH), 1745 (ester CO), 1719 (carbamate CO), 1652 (amide CO) and 736 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.27 (_c_, 3H, d, _J_ 7.0, CHC$\underline{H}_3$), 1.33 (_t_, 3H, d, _J_ 6.6, CHC$\underline{H}_3$), 1.79-2.33 (_t_ & _c_, 4H, m, βCH_2 and γCH_2), 3.38-3.62 (_t_ & _c_, 2H, m, δCH_2), 3.70 (_t_, 3H, s, OCH_3), 3.74 (_c_, 3H, s, OCH_3), 4.31 (_c_, 1H, dq, J_1 7.0, J_2 8.0, C$\underline{H}$CH$_3$), 4.42 (_t_, 1H, dd, J_1 2.1, J_2 9.1, αCH), 4.54 (_t_, 1H, dq, J_1 6.6, J_2 7.6, C$\underline{H}$CH$_3$), 4.91 (_c_, 1H, dd, J_1 3.1, J_2 7.3, αCH), 5.08 (_t_ & _c_, 2H, s, CH_2Ph), 5.59 (_c_, 1H, d, _J_ 8.0, NH), 5.78 (_t_, 1H, d, _J_ 7.6, NH), 7.31 (_c_, 5H, s, aromatic) and 7.32 (_t_, 5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 17.89 (_c_, CH$\underline{C}$H$_3$), 18.48 (_t_, CH$\underline{C}$H$_3$), 22.67 (_c_, γCH_2), 24.92 (_t_, γCH_2), 29.22 (_t_, βCH_2), 31.31 (_c_, βCH_2), 46.81 (_c_, δCH_2), 47.00 (_t_, δCH_2), 48.42 (_c_, $\underline{C}$HCH$_3$), 48.67 (_t_, $\underline{C}$HCH$_3$), 52.25 (_t_, OCH_3), 52.72 (_c_, OCH_3), 59.32 (_t_, αCH), 60.46 (_c_, αCH), 66.63 (_t_ & _c_, CH_2Ph), 128.13 (_t_ & _c_, aromatic), 128.64 (_t_ & _c_, aromatic), 136.93 (_c_, quat. aromatic), 137.03 (_t_, quat. aromatic), 155.77 (_t_, $\underline{C}O_2CH_2Ph$), 156.37 (_c_, $\underline{C}O_2CH_2Ph$), 171.05 (_t_, $\underline{C}OCHCH_3$), 171.29 (_c_, $\underline{C}OCHCH_3$), 172.53 (_t_, $\underline{C}O_2CH_3$) and 172.98 (_t_, $\underline{C}O_2CH_3$); _m/z_ (CI) 335 ($[M + H]^+$, 100%), 230 (5, $[M - CH_2Ph - CH_3 + 2H]^+$), 201 (3, $[M - CO_2CH_2Ph + 2H]^+$), 196 (4, $[M - OCH_2Ph - OCH_3]^+$) and 169 (6, $[M - CO_2CH_2Ph - OCH_3 + H]^+$).

_N_-Benzyloxycarbonyl-glycyl-(2_S_)-prolinamide (60)

To a saturated solution of ammonia in dry methanol (20 cm^3) was added _N_-benzyloxycarbonyl-glycyl-(2_S_)-proline methyl ester (**57**) (1.60 g, 5 mmol). The solution was placed in a tightly stoppered vessel and left at room temperature. When the reaction was complete as judged by TLC (9 days), the stoppered vessel was cooled to 0 °C, opened and nitrogen gas bubbled through at room temperature for 5 min. The solvent was then removed under reduced pressure to give a white solid which was recrystallised from methanol/ diethyl ether to give the product as a white solid (1.24 g, 81%), m.p. 148-9 °C; (Found: C, 59.15; H, 6.4; N, 13.8. Calc. for $C_{15}H_{19}N_3O_4$: C, 59.0; H, 6.3; N, 13.75%); _m/z_ (Found: $[M + H]^+$, 306.1459. Calc. for $C_{15}H_{20}N_3O_4$: 306.1454); $[\alpha]_D^{22}$ -60.1 (_c_ 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3394 &

3213 (NH), 1687 (carbamate CO), 1671 (tertiary amide CO), 1639 (primary amide CO) and 753 & 701 (aromatic CH); δ_H (200 MHz; d_6-DMSO) 1.67-2.34 (*t* & *c*, 4H, m, γCH_2 and βCH_2), 3.31-3.69 (*t* & *c*, 2H, m, δCH_2), 3.79 (*c*, 2H, d, *J* 5.8, $COCH_2$), 3.86 (*t*, 2H, d, *J* 5.8, $COCH_2$), 4.21 (*t*, 1H, dd, J_1 5.8, J_2 7.6, αCH), 4.35 (*c*, 1H, dd, J_1 2.6, J_2 8.2, αCH), 5.01 (*t* & *c*, 2H, s, CH_2Ph), 7.01 (*t*, 1H, s, $\frac{1}{2}CONH_2$), 7.28 (*c*, 1H, s, $\frac{1}{2}CONH_2$), 7.32 (*t*, 1H, s, $\frac{1}{2}CONH_2$), 7.38 (*t* & *c*, 5H, s, aromatic) and 7.61 (*c*, 1H, s, $\frac{1}{2}CONH_2$); δ_C (50.31 MHz; $C^2H_3O^2H$) 23.62 (*c*, γCH_2), 25.90 (*t*, γCH_2), 31.05 (*t*, βCH_2), 33.69 (*c*, βCH_2), 44.35 (*t*, δCH_2), 44.46 (*c*, δCH_2), 47.83 (*t*, $COCH_2$), 48.61 (*c*, $COCH_2$), 61.26 (*c*, αCH), 61.86 (*t*, αCH), 68.01 (*t* & *c*, CH_2Ph), 129.16 (*t* & *c*, *para* aromatic), 129.28 (*t* & *c*, *ortho* aromatic), 129.74 (*t* & *c*, *meta* aromatic), 138.43 (*t* & *c*, quat. aromatic), 159.30 (*t* & *c*, $\underline{C}O_2CH_2Ph$), 170.58 (*t*, $\underline{C}OCH_2$) , 170.72 (*c*, $\underline{C}OCH_2$), 177.49 (*c*, $CONH_2$) and 174.94 (*t*, $CONH_2$); *m/z* (CI) 306 ([*M* + H]$^+$, 33%), 198 (100, [*M* - OCH_2Ph]$^+$), 172 (11, [*M* - CO_2CH_2Ph + 2H]$^+$), 155 (38, [*M* - CO_2CH_2Ph - NH_2 + H]$^+$), 139 (15, [*M* - CO_2CH_2Ph - NH_2 - NH]$^+$) and 91 (22, [CH_2Ph]$^+$).

_N_-Benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (61)



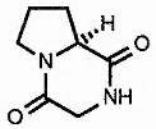
This compound was prepared in a manner identical with that for *N*-benzyloxycarbonyl-glycyl-(2*S*)-prolinamide (**60**), starting from the *N*-*tert*-butoxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (**58**) (0.67 g, 2 mmol) to give an off-white solid which was recrystallised from methanol/ diethyl ether to give the product as colourless crystals (0.51 g, 79%), m.p. 167-9 °C; (Found: C, 60.5; H, 6.75; N, 13.3. Calc. for $C_{16}H_{21}N_3O_4$: C, 60.2; H, 6.65; N, 13.2%); *m/z* (Found: [*M* + H]$^+$, 320.1612. Calc. for $C_{16}H_{22}N_3O_4$: 320.1610); $[\alpha]_D^{22}$ -81.3 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3402 & 3323 (NH), 1668 (carbamate CO), 1653 (tertiary amide CO), 1646 (primary amide CO) and 754 & 696 (aromatic CH); δ_H (200 MHz; d_6-DMSO) 1.20 (3H, d, *J* 7.6, CH_3), 1.62-2.18 (4H, m, βCH_2 and γCH_2), 3.58 (2H, m, δCH_2), 4.22 (1H, dd, J_1 7.5, J_2 2.7, αCH), 4.32 (1H, m, $C\underline{H}CH_3$), 5.01 (2H, s, CH_2Ph), 6.88 (1H, s,$\frac{1}{2}CONH_2$), 7.22 (1H, s, $\frac{1}{2}CONH_2$), 7.35 (5H, s, aromatic) and 7.51 (1H, d, *J* 7.6, NH); δ_C (74.76 MHz; $C^2H_3O^2H$) 17.63 (*t*, CH_3), 18.28 (*c*, CH_3), 23.52 (*c*, γCH_2), 26.42 (*t*, γCH_2), 31.17 (*t*, βCH_2), 33.36 (*c*, βCH_2), 48.93 (*t* & *c*,

δCH_2), 50.36 (*t*, $\underline{C}HCH_3$), 50.72 (*c*, $\underline{C}HCH_3$), 61.86 (*t*, αCH), 62.05 (*c*, αCH), 68.11 (*t*, CH_2Ph), 68.73 (*c*, CH_2Ph), 129.32 (*t* & *c*, *ortho* aromatic), 129.49 (*t* & *c*, *para* aromatic), 129.94 (*t* & *c*, *meta* aromatic), 138.66 (*t* & *c*, quat. aromatic), 158.66 (*t* & *c*, $\underline{C}O_2CH_2Ph$), 174.45 (*t* & *c*, $\underline{C}OCHCH_3$), 176.93 (*c*, $CONH_2$) and 177.54 (*t*, $CONH_2$); *m/z* (CI) 320 ([*M* + H]$^+$, 100%), 303 (20, [*M* - NH_2]$^+$), 291 (21, [*M* - CO]$^+$), 276 (19, [*M* - $CONH_2$ + H]$^+$), 227 (34, [*M* - $PhCH_2$ + H]$^+$), 212 (38, [*M* - $PhCH_2O$]$^+$), 91 (40, [$PhCH_2$]$^+$) and 70 (23, [C_4H_8N]$^+$).

_N_-Benzyloxycarbonyl-(2*R*)-alanyl-(2*S*)-prolinamide (62)

This compound was prepared in a manner identical with that for the *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (**61**), using *N*-benzyloxycarbonyl-(2*R*)-alanyl-(2*S*)-proline methyl ester (**59**) (1.60 g, 5 mmol) instead of *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (**58**) to give the product as a clear oil (1.57 g, 98%), *m/z* (Found: [*M* + H]$^+$, 320.1606. Calc. for $C_{16}H_{22}N_3O_4$: 320.1610); $[\alpha]_D^{22}$ -19.8 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 1701 (carbamate CO), 1683 (tertiary amide CO), 1642 (primary amide CO) and 743 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.30 (3H, d, *J* 6.8, $CHC\underline{H}_3$), 1.70-2.43 (4H, m, βCH_2 and γCH_2), 3.22-3.60 (1H, m, $\frac{1}{2}\delta CH_2$), 3.62-3.97 (1H, m, $\frac{1}{2}\delta CH_2$), 4.23-4.61 (2H, m, αCH and $C\underline{H}CH_3$), 5.04 (2H, s, CH_2Ph), 5.79 (1H, s, $\frac{1}{2}NH_2$), 6.10 (1H, d, *J* 6.6, $N\underline{H}CHCH_3$), 6.86 (1H, s, $\frac{1}{2}NH_2$) and 7.32 (5H, s, aromatic); δ_C (50.31 MHz; $C^2H_3O^2H$) 17.44 (*t*, CH_3), 18.12 (*c*, CH_3), 23.96 (*c*, γCH_2), 25.89 (*t*, γCH_2), 31.08 (*t*, βCH_2), 33.77 (*c*, βCH_2), 48.42 (*c*, δCH_2), 48.75 (*t*, δCH_2), 50.04 (*t* & *c*, $\underline{C}HCH_3$), 62.03 (*c*, αCH), 62.29 (*t*, αCH), 67.87 (*c*, CH_2Ph), 68.13 (*t*, CH_2Ph), 129.24 (*t* & *c*, *para* aromatic), 129.37 (*t* & *c*, *ortho* aromatic), 129.95 (*t* & *c*, *meta* aromatic), 138.46 (*t* & *c*, quat. aromatic), 158.56 (*c*, $\underline{C}O_2CH_2Ph$), 158.68 (*t*, $\underline{C}O_2CH_2Ph$), 174.36 (*t*, $\underline{C}OCHCH_3$), 175.02 (*c*, $\underline{C}OCHCH_3$), 177.24 (*c*, CO_2NH_2) and 177.39 (*t*, CO_2NH_2); *m/z* (CI) 320 ([*M* + H]$^+$, 100%), 303 (32, [*M* - NH_2]$^+$), 291 (15, [*M* - CO]$^+$), 276 (22, [*M* - $CONH_2$ + H]$^+$), 230 (31, [*M* - Ph - CH_3 + 3H]$^+$), 186 (36, [*M* - CO_2CH_2Ph + 2H]$^+$), 91 (52, [$PhCH_2$]$^+$) and 70 (18, [C_4H_8N]$^+$).

(8a*S*)-2,3,6,7,8,8a-Hexahydropyrrolo-[1,2-a]-pyrazine-1,4-dione (64)



To a solution of *N*-benzyloxycarbonyl-glycyl-(2*S*)-prolinamide (**60**) (0.610 g, 2 mmol) in methanol (50 cm^3) was added 10% palladium on activated charcoal (30 mg) and the vessel flushed with hydrogen. The resultant suspension was stirred vigorously under an atmosphere of hydrogen for 12 h. The mixture was filtered through a celite pad and the solvent removed under reduced pressure to give white solid which was recrystallised from methanol to give the product as colourless crystals (0.29 g, 93%), m.p. 205-8 °C (lit.,[240] 208-10 °C); (Found: C, 54.45; H, 6.8; N, 18.4; M^+, 154.0747. Calc. for $C_7H_{10}N_2O_2$: C, 54.5; H, 6.65; N, 18.3%; M^+, 154.0742); $[\alpha]_D^{22}$ -184.3 (*c* 0.5 in MeOH) {lit.,[240] -196.5 (*c* 0.5 in MeOH)}; υ_{max} (Nujol)/cm^{-1} 3162 & 3110 (NH), 1681 (tertiary amide CO) and 1652 (secondary amide CO); δ_H (300 MHz; $C^2H_3O^2H$) 2.03-2.38 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.45-2.64 (1H, m, $\frac{1}{2}\beta CH_2$), 3.73 (2H, m, δCH_2), 3.93 (1H, d, *J* 16.8, $\frac{1}{2}COCH_2$), 4.30 (1H, d, *J* 16.8, $\frac{1}{2}COCH_2$) and 4.42 (1H, dd, J_1 6.8, J_2 6.8, αCH); δ_C (74.76 MHz; $C^2H_3O^2H$) 23.60 (γCH_2), 29.68 (βCH_2), 46.60 ($CO\underline{C}H_2$), 47.29 (δCH_2), 60.15 (αCH), 166.74 ($\underline{C}OCH_2$) and 172.27 (CONH); *m/z* (EI) 154 (M^+, 70%), 126 (8, [*M* - $NHCH_2$ + H]$^+$), 111 (81, [*M* - CONH]$^+$), 98 (28, [*M* - $CONHCH_2$ + 2H]$^+$), 83 (100, [*M* - CO_2 - $NHCH_2$ + 2H]$^+$) and 70 (78, $[C_4H_8N]^+$).

***N-tert*-Butoxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (66)**

This compound was prepared in a manner identical with that for the *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (**58**), starting from *N-tert*-butoxycarbonyl-(2*S*)-alanine (**65**) (1.89 g, 10 mmol) to give a yellow oil which was purified by silica chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a clear oil (2.37 g, 79%), *m/z* (Found: [*M* + H]$^+$, 301.1769. Calc. for $C_{14}H_{25}N_2O_5$: 301.1764); $[\alpha]_D^{22}$ -109.4 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3325 (NH), 1748 (ester CO), 1710 (carbamate CO) and 1652 (amide CO); δ_H (200 MHz; C^2HCl_3) 1.26 (*c*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.33 (*t*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.40 (*t* & *c*, 9H, s, $C(CH_3)_3$), 1.82-2.33 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.50-3.76 (*t* & *c*, 2H, m, δCH_2), 3.70 (*t*, 3H, s, OCH_3), 3.74 (*c*, 3H, s, OCH_3), 4.48 (*t* & *c*, 2H, m, αCH and $C\underline{H}CH_3$) and 5.34 (*t* & *c*, 1H, d, *J* 7.9, NH);

δ_C (50.31 MHz; C^2HCl_3) 17.86 (*t*, CH<u>C</u>H$_3$), 18.96 (*c*, CH<u>C</u>H$_3$), 21.94 (*c*, γCH$_2$), 24.83 (*t*, γCH$_2$), 28.20 (*t* & *c*, C(<u>C</u>H$_3$)$_3$), 28.78 (*t*, βCH$_2$), 31.14 (*c*, βCH$_2$), 46.26 (*c*, δCH$_2$), 46.60 (*t*, δCH$_2$), 47.60 (*t*, OCH$_3$), 47.73 (*c*, OCH$_3$), 51.92 (*t*, C<u>H</u>CH$_3$), 52.52 (*c*, C<u>H</u>CH$_3$), 58.57 (*t*, αCH), 58.95 (*c*, αCH), 125.88 (*t* & *c*, <u>C</u>(CH$_3$)$_3$), 154.58 (*c*, <u>C</u>O$_2$^{t}Bu), 155.06 (*t*, <u>C</u>O$_2$^{t}Bu), 171.44 (*t* & *c*, <u>C</u>OCHCH$_3$), 172.08 (*c*, <u>C</u>O$_2$CH$_3$) and 172.23 (*t*, <u>C</u>O$_2$CH$_3$); *m/z* (CI) 301 ([*M* + H]$^+$, 52%), 245 (100, [*M* - C(CH$_3$)$_3$ + 2H]$^+$), 227 (70, [*M* - OC(CH$_3$)$_3$]$^+$), 201 (17, [*M* - CO$_2$C(CH$_3$)$_3$ + 2H]$^+$), 169 (36, [*M* - CO$_2$C(CH$_3$)$_3$ - OCH$_3$ + H]$^+$) and 153 (6, [*M* - NHCO$_2$C(CH$_3$)$_3$ - OCH$_3$]$^+$).

N-tert-Butoxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (67)

This compound was prepared in a manner identical with that for *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (**61**), starting from the *N-tert*-butoxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (**66**) (1.50 g, 5 mmol) to give an off-white solid which was recrystallised from methanol/ diethyl ether to give the product as a white solid (1.11 g, 78%), m.p. 168-70 °C; *m/z* (Found: [*M* + H]$^+$, 286.1759. Calc. for $C_{13}H_{25}N_3O_4$: 287.1764); $[\alpha]_D^{22}$ -81.4 (*c* 0.5 in MeOH); υ_{max} (Nujol)/cm^{-1} 3404 & 3334 (NH), 1701 (carbamate CO), 1670 (tertiary amide CO) and 1648 (primary amide CO); δ_H (200 MHz; C^2HCl_3) 1.28 (*c*, 3H, d, *J* 7.0, CHC<u>H</u>$_3$), 1.33 (*t*, 3H, d, *J* 6.8, CHC<u>H</u>$_3$), 1.71-2.23 (*t* & *c*, 3H, m, γCH$_2$ and $\frac{1}{2}\beta$CH$_2$), 2.28-2.45 (*t* & *c*, 1H, m, $\frac{1}{2}\beta$CH$_2$), 3.48-3.79 (*t* & *c*, 2H, m, δCH$_2$), 4.15 (*c*, 1H, m, C<u>H</u>CH$_3$), 4.21 (*c*, 1H, dd, J_1 1.9, J_2 8.0, αCH), 4.48 (*t*, 1H, m, C<u>H</u>CH$_3$), 4.61 (*t*, 1H, dd, J_1 2.6, J_2 7.8, αCH), 5.20 (*c*, 1H, d, *J* 7.4, N<u>H</u>CHCH$_3$), 5.35 (*t*, 1H, d, *J* 7.2, N<u>H</u>CHCH$_3$), 5.54 (*t*, 1H, s, $\frac{1}{2}$CONH$_2$), 5.79 (*c*, 1H, s, $\frac{1}{2}$CONH$_2$), 6.82 (*t*, 1H, s, $\frac{1}{2}$CONH$_2$) and 7.43 (*c*, 1H, s, $\frac{1}{2}$CONH$_2$); δ_C (50.31 MHz; C^2HCl_3) 18.29 (*t*, CH<u>C</u>H$_3$), 18.42 (*c*, CH<u>C</u>H$_3$), 22.54 (*c*, γCH$_2$), 25.52 (*t*, γCH$_2$), 28.84 (*t* & *c*, C(<u>C</u>H$_3$)$_3$), 28.88 (*t*, βCH$_2$), 32.13 (*c*, βCH$_2$), 47.33 (*c*, δCH$_2$), 47.73 (*t*, δCH$_2$), 48.39 (*t*, C<u>H</u>CH$_3$), 49.00 (*c*, C<u>H</u>CH$_3$), 60.23 (*t* & *c*, αCH), 79.99 (*t*, <u>C</u>(CH$_3$)$_3$), 80.38 (*c*, <u>C</u>(CH$_3$)$_3$), 155.90 (*t* & *c*, CO$_2$^{t}Bu), 173 31 (*t* & *c*, <u>C</u>OCHCH$_3$) and 174 94 (*t* & *c*, CONH$_2$); *m/z* (CI) 286 ([*M* + H]$^+$, 100%), 230 (52, [*M* - C(CH$_3$)$_3$ + 2H]$^+$), 212 (3, [*M* - OC(CH$_3$)$_3$]$^+$), 186 (5, [*M* - CO$_2$C(CH$_3$)$_3$ + 2H]$^+$) and 143 (3, [*M* - CO$_2$C(CH$_3$)$_3$ - CONH$_2$ + 3H]$^+$).

Glycyl-(2*S*)-prolinamide hydrochloride (68)

Hydrogen chloride gas was bubbled through a solution of *N-tert*-butoxycarbonyl-glycyl-(2*S*)-prolinamide (**23**) (0.81 g, 3 mmol) in ethylacetate (30 cm^3) for 20 min at

0 °C. The solvent was then removed under reduced pressure to give the product as a hygroscopic white solid (0.60 g, 97%), m.p. 143-5 °C; *m/z* (Found: M^+, 172.2063. Calc. for $C_7H_{14}N_3O_2$: 172.2065); $[\alpha]_D^{22}$ -52.5 (*c* 0.2 in MeOH); υ_{max} (Nujol)/cm^{-1} 1662 (tertiary amide CO) and 1652 (primary amide CO); δ_H (200 MHz; $C^2H_3O^2H$) 1.73-2.47 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.42-3.76 (*t* & *c*, 2H, m, δCH_2), 3.93 (*t* & *c*, 2H, s, $COCH_2$), 4.49 (*t*, 1H, dd, J_1 7.5, J_2 11.8, αCH) and 4.56 (*c*, 1H, dd, J_1 4.1, J_2 15.2, αCH); δ_C (50.31 MHz; $C^2H_3O^2H$) 23.91 (*c*, γCH_2), 26.03 (*t*, γCH_2), 31.48 (*t*, βCH_2), 33.83 (*c*, βCH_2), 42.07 (*c*, $CO\underline{C}H_2$), 42.48 (*t*, $CO\underline{C}H_2$), 48.40 (*t*, δCH_2), 48.96 (*c*, δCH_2), 61.29 (*c*, αCH), 61.74 (*t*, αCH), 167.08 (*c*, $CONH_2$), 167.43 (*t*, $CONH_2$), 176.68 (*c*, $\underline{C}OCH_2$) and 178.51 (*t*, $\underline{C}OCH_2$); *m/z* (CI) 172 (M^+, 15%), 155 (100, $[M - NH_3]^+$), 111 (8, $[M - NH_3 - CONH_2]^+$) and 70 (11, $[C_4H_8N]^+$).

(2*S*)-Alanyl-(2*S*)-prolinamide hydrochloride (69)

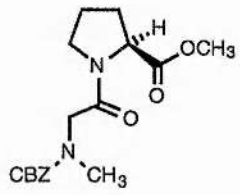
This compound was prepared in a manner identical with that for the glycyl-(2*S*)-prolinamide hydrochloride (**68**), using *N-tert*-butoxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (**67**) (2.43 g, 5 mmol) instead of *N-tert*-butoxycarbonyl-glycyl-(2*S*)-prolinamide (**23**) to give the product as a hygroscopic white solid (1.06 g, 96%), *m/z* (Found: $[M + H]^+$, 186.1244. Calc. for $C_8H_{16}N_3O_2$: 186.1242); $[\alpha]_D^{22}$ -9.8 (*c* 0.1 in MeOH); υ_{max} (Nujol)/cm^{-1} 1668 (tertiary amide CO) and 1652 (primary amide CO); δ_H (200 MHz; ^2H_2O) 1.38 (*c*, 3H, d, *J* 7.0, $CHC\underline{H}_3$), 1.43 (*t*, 3H, d, *J* 7.0, $CHC\underline{H}_3$), 1.72-2.08 (*t* & *c*, 3H, m, $\frac{1}{2}\beta CH_2$ and γCH_2), 2.24 (*t* & *c*, 1H, m, $\frac{1}{2}\beta CH_2$), 3.42-3.71 (*t* & *c*, 1H, q, *J* 7.0, $C\underline{H}CH_3$) and 4.34 (*t* & *c*, 1H, dd, J_1 6.8, J_2 8.6, αCH); δ_C (50.31 MHz; ^2H_2O) 15.86 (*t*, $CH\underline{C}H_3$), 16.44 (*c*, $CH\underline{C}H_3$), 22.64 (*c*, γCH_2), 25.56 (*t*, γCH_2), 30.38 (*t*, βCH_2), 32.78 (*c*, βCH_2), 48.41 (*c*, δCH_2), 48.56 (*t*, δCH_2), 48.66 (*t*, $\underline{C}HCH_3$), 48.88 (*c*, $\underline{C}HCH_3$), 60.57 (*c*, αCH), 61.17 (*t*, αCH), 170.07 (*t* & *c*, $\underline{C}OCH_3$), 176.82 (*c*, $CONH_2$) and 177.30 (*t*, $CONH_2$); *m/z* (EI) 186 ($[M + H]^+$, 4%), 168 (88, $[M - NH_3]^+$), 140 (14, $[M - NH_3 - CO]^+$), 125 (56, $[M - NH_3 - CH_3 - CO]^+$), 97 (69, $[M - NH_3 - CH_3 - 2CO]^+$) and 70 (100, $[C_4H_8N]^+$).

***N*-Benzyloxycarbonyl-sarcosine (71)**

To a vigorously stirred solution of sarcosine (**70**) (2.05 g, 23 mmol) in 1 mol dm^{-3} NaOH (80 cm^3, 80 mmol) was added dropwise benzylchloroformate (5.4 cm^3, 37 mmol) at 0 °C. The resultant cloudy solution was stirred vigorously for 4 h at room temperature and then washed with diethyl ether (50 cm^3). The aqueous phase

was acidified to pH 3 with 1 mol dm^{-3} HCl and then extracted with diethyl ether (3 x 50 cm^3). The ethereal extracts were washed with 1 mol dm^{-3} HCl (50 cm^3) and saturated NaCl solution (50 cm^3), dried($MgSO_4$), and the solvent removed under reduced pressure to give the product as a clear oil (2.14 g, 96%), *m/z* (Found: M^+, 223.0853. Calc. for $C_{11}H_{13}NO_4$: 223.0845); υ_{max} (Nujol)/cm^{-1} 1701 (acid CO) and 1662 (carbamate CO); δ_H (200 MHz; C^2HCl_3) 3.03 (*t* & *c*, 3H, s, NCH_3), 4.04 (*c*, 2H, s, $COCH_2$), 4.10 (*t*, 2H, s, $COCH_2$), 5.15 (*c*, 2H, s, CH_2Ph), 5.17 (*t*, 2H, s, CH_2Ph), 7.32 (*c*, 5H, s, aromatic), 7.37 (*t*, 5H, s, aromatic) and 10.02 (*t* & *c*, 1H, s, CO_2H); δ_C (50.31 MHz; C^2HCl_3) 35.76 (*c*, NCH_3) 36.41 (*t*, NCH_3), 50.64 (*c*, $CO\underline{C}H_2$), 50.91 (*t*, $CO\underline{C}H_2$), 68.23 (*c*, CH_2Ph), 68.30 (*t*, CH_2Ph), 128.56 (*c*, *para* aromatic), 128.63 (*t*, *para* aromatic), 128.24 (*c*, *ortho* aromatic), 128.33 (*t*, *ortho* aromatic), 128.91 (*c*, *meta* aromatic), 129.02 (*t*, *meta* aromatic), 136.74 (*t* & *c*, quat. aromatic), 157.05 (*c*, $\underline{C}ONCH_3$), 157.70 (*t*, $\underline{C}ONCH_3$), 173.46 (*c*, CO_2H) and 173.62 (*t*, CO_2H); *m/z* (EI) 223 (M^+, 6%), 134 (4, [*M* - $PhCH_2$ + 2H]$^+$), 108 (39, [$PhCH_2OH$]$^+$) and 91 (100, [$PhCH_2$]$^+$).

_N_-Benzyloxycarbonyl-sarcosyl-(2*S*)-proline methyl ester (72)



This compound was prepared in a manner identical with that for the *N* benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (**58**), starting from *N* benzyloxycarbonyl-sarcosine (**71**) (2.23 g, 10 mmol) to give a yellow oil which was purified by silica column chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a clear oil (2.56 g, 77%), *m/z* (Found: $[M + H]^+$, 335.1599. Calc. for $C_{17}H_{23}N_2O_5$: 335.1606); $[\alpha]_D^{22}$ -78.3 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 2954 (N-CH_3), 1747 (ester CO), 1710 (carbamate CO), 1663 (amide CO) and 734 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.68-2.41 (*tt*, *tc*, *ct* & *cc*, 4H, m, γCH_2 and βCH_2), 2.96 (*cc*, 3H, s, NCH_3), 2.97 (*tt* & *tc*, 3H, s, NCH_3), 2.99 (*ct*, 3H, s, NCH_3), 3.22-3.39 (*cc*, 2H, m, δCH_2), 3.39-3.75 (*tt*, *ct* & *tc*, 2H, m, δCH_2), 3.65 (*cc*, 3H, s, OCH_3), 3.66 (*tt*, 3H, s, OCH_3), 3.68 (*ct*, 3H, s, OCH_3), 3.69 (*tc*, 3H, s, OCH_3), 3.88 (*tt*, 1H, d, *J* 16.5, $\frac{1}{2}COCH_2$), 3.90 (*ct*, 1H, d, *J* 16.5, $\frac{1}{2}COCH_2$), 3.97 (*tc*, 2H, d, *J* 5.9, $COCH_2$), 3.99 (*cc*, 2H, d, *J* 6.0, $COCH_2$), 4.09 (*cc*, 1H, dd, $J_1 = J_2$ 3.2, αCH), 4.21 (*tt*, 1H, d, *J* 16.5, $\frac{1}{2}COCH_2$), 4.22 (*ct*, 1H, d, *J* 16.5,

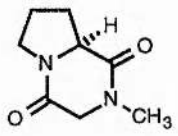
$\frac{1}{2}COCH_2$), 4.41 (*tc*, 1H, dd, $J_1 = J_2$ 2.9, αCH), 4.47 (*tt* & *ct*, dd, J_1 3.6, J_2 5.5, αCH), 5.07 (*cc*, 2H, s, CH_2Ph), 5.09 (*tt* & *tc*, 2H, s, CH_2Ph), 5.10 (*ct*, 2H, s, CH_2Ph), 7.28 (*cc*, 5H, s, aromatic), 5.09 (*tt* & *tc*, 5H, s, aromatic) and 5.10 (*ct*, 5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 22.36 (*tc* & *cc*, γCH_2), 25.05 (*ct*, γCH_2), 25.12 (*tt*, γCH_2), 29.09 (*tt* & *ct*, βCH_2), 31.48 (*cc*, βCH_2), 31.58 (*tc*, βCH_2), 35.48 (*tt*, NCH_3), 35.61 (*tc* & *cc*, NCH_3), 36.28 (*ct*, NCH_3), 46.27 (*tt* & *ct*, δCH_2), 46.81 (*tc* & *cc*, δCH_2), 50.99 (*ct*, $CO\underline{C}H_2$), 51.28 (*tt*, $CO\underline{C}H_2$), 51.42 (*tc* & *cc*, $CO\underline{C}H_2$), 52.28 (*tt* & *ct*, OCH_3), 52.74 (*cc*, OCH_3), 52.82 (*tc*, OCH_3), 58.70 (*cc*, αCH), 58.79 (*tc*, αCH), 59.15 (*tt* & *ct*, αCH), 67.23 (*ct* & *cc*, CH_2Ph), 67.34 (*tt* & *tc*, CH_2Ph), 127.88 (*tt*, *tc*, *ct* & *cc*, *para* aromatic), 128.08 (*ct* & *cc*, *meta* aromatic), 128.14 (*tt* & *tc*, *meta* aromatic), 128.40 (*tc*, *ortho* aromatic), 128.69 (*tt*, *ct* & *cc*, *ortho* aromatic), 137.10 (*ct* & *cc*, quat. aromatic), 137.16 (*tt* & *tc*, quat. aromatic), 156.44 (*ct* & *cc*, $\underline{C}O_2CH_2Ph$), 156.92 (*tt* & *tc*, $\underline{C}O_2CH_2Ph$), 167.34 (*tt* & *ct*, $\underline{C}OCH_2$), 167.69 (*tc* & *cc*, $\underline{C}OCH_2$), 172.43 (*cc*, $\underline{C}O_2CH_3$), 172.58 (*tc*, $\underline{C}O_2CH_3$), 172.66 (*ct*, $\underline{C}O_2CH_3$) and 172.74 (*tt*, $\underline{C}O_2CH_3$); *m/z* (CI) 335 ([*M* + H]$^+$, 100%), 291 (12, [*M* - CH_3 - OCH_3 + 3H]$^+$), 226 (5, [*M* - OCH_3 - Ph]$^+$), 201 (6, [*M* - CO_2CH_2Ph + 2H]$^+$) and 91 (7, [$PhCH_2$]$^+$).

_N_-Benzyloxycarbonyl-sarcosyl-(2_S_)-prolinamide (73)

This compound was prepared in a manner identical with that for the *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (**61**), starting from *N*-benzyloxycarbonyl-sarcosyl-(2*S*)-proline methyl ester (**72**) (1.67 g, 5 mmol) to give the product as a clear oil (1.42 g, 80%), *m/z* (Found: [*M* + H]$^+$, 320.1613. Calc. for $C_{16}H_{22}N_3O_4$: 320.1610); $[\alpha]_D^{22}$ +3.0 (*c* 0.1 in MeOH); υ_{max} (thin film)/cm^{-1} 3326 & 3195 (NH), 1701 (carbamate CO), 1668 (tertiary amide CO), 1652 (primary amide CO) and 737 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.66-2.32 (*tt* & *ct*, 4H, m, βCH_2 and γCH_2), 3.06 (*ct*, 3H, s, NCH_3), 3.07 (*tt*, 3H, s, NCH_3), 3.20-3.73 (*tt* & *ct*, 2H, m, $COCH_2$), 4.01 (*ct*, 2H, m, δCH_2), 4.04 (*tt*, 2H, m, δCH_2), 4.58 (*tt* & *ct*, 1H, dd, J_1 2.6, J_2 8.3, αCH), 5.15 (*tt* & *ct*, 2H, s, CH_2Ph), 5.91 (*ct*, 1H, s, $\frac{1}{2}CONH_2$), 6.00 (*tt*, 1H, s, $\frac{1}{2}CONH_2$), 6.92 (*ct*, 1H, s, $\frac{1}{2}CONH_2$), 6.99 (*tt*, 1H, s, $\frac{1}{2}CONH_2$), 7.32 (*ct*, 5H, s, aromatic) and 7.36 (*tt*, 5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 22.41 (*tc* & *cc*, γCH_2), 24.84 (*tt* & *ct*, γCH_2), 28.76 (*ct*, βCH_2), 29.03 (*tt*, βCH_2), 32.33 (*tc* & *cc*, βCH_2), 35.77 (*tt* & *tc*, NCH_3), 36.32 (*ct* & *cc*, NCH_3), 46.51 (*tt* & *ct*, δCH_2), 47.15 (*tc* & *cc*, δCH_2), 50.91 (*ct* & *cc*, $CO\underline{C}H_2$), 51.37 (*tt* & *tc*, $CO\underline{C}H_2$), 60.05 (*ct* & *cc*, αCH), 60.18 (*tt* & *tc*, αCH), 67.25 (*ct* & *cc*, CH_2Ph), 67.35 (*tt* & *tc*, CH_2Ph), 127.60 (*ct* & *cc*, *para* aromatic), 127.69 (*tt* & *tc*, *para* aromatic), 127.81 (*ct* & *cc*, *ortho* aromatic), 128.03 (*tt* & *tc*, *ortho* aromatic), 128.52 (*tt*, *tc*, *ct* & *cc*, *meta* aromatic),

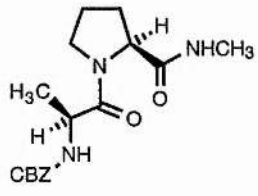
136.64 (*tt* & *tc*, quat. aromatic), 136.69 (*ct* & *cc*, quat. aromatic), 156.38 (*ct* & *cc*, quat. aromatic), 156.38 (*ct* & *cc*, $\underline{C}O_2CH_2Ph$), 157.01 (*tt* & *tc*, $\underline{C}O_2CH_2Ph$), 167.96 (*tt* & *tc*, $N\underline{C}O_2Ph$), 168.07 (*ct* & *cc*, $N\underline{C}O_2Ph$), 174.53 (*ct* & *cc*, $CONH_2$) and 174.80 (*tt* & *tc*, $CONH_2$); *m/z* (CI) 320 ([*M* + H]$^+$, 42%), 186 (6, [*M* - CO_2CH_2Ph + 2H]$^+$), 113 (93, [*M* - CO_2CH_2Ph - $COCH_2NCH_3$]$^+$), 97 ([100, $[C_4H_8NCO]^+$) and 70 (70, $[C_4H_8N]^+$).

(8a*S*)-2-Methyl-2,3,6,7,8,8a-hexahydropyrrolo-[1,2-a]-pyrazine-1,4-dione (74)



This compound was prepared in a manner identical with that for the hexahydropyrrolo-[1,2-a]-pyrazine-1,4-dione (**64**), using *N*-benzyloxycarbonyl-sarcosyl-(2*S*)-prolinamide (**73**) (0.64 g, 2 mmol) to give the product as a clear oil (0.31 g, 90%), *m/z* (Found: *M*$^+$, 168.0901. Calc. for $C_8H_{12}N_2O_2$: 168.0899); $[\alpha]_D^{22}$ +4.9 (*c* 0.1 in MeOH); υ_{max} (thin film)/cm^{-1} 1675 (tertiary amide CO) and 1652 (secondary amide CO); δ_H (200 MHz; $C^2H_3O^2H$) 1.81-2.12 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.13-2.49 (1H, m, $\frac{1}{2}\beta CH_2$), 2.99 (3H, s, NCH_3), 3.42-3.73 (2H, m, δCH_2), 3.83 (1H, d, *J* 16.8, $\frac{1}{2}COCH_2$), 4.23 (1H, m, αCH) and 4.30 (1H, d, *J* 16.8, $\frac{1}{2}COCH_2$); δ_C (50.31 MHz $C^2H_3O^2H$) 23.64 (γCH_2), 30.21 (βCH_2), 33.95 (NCH_3), 46.54 (δCH_2), 54.40 ($CO\underline{C}H_2$), 60.38 (αCH), 165.62 ($\underline{C}ONCH_3$) and 169.74 ($\underline{C}OCH_2$); *m/z* (EI) 168 (*M*$^+$, 45%), 156 (18, [*M* - CH_3 + 3H]$^+$), 140 (15, [*M* - NCH_3 + H]$^+$), 127 (12, [*M* - CH_2NCH_3 + 2H]$^+$), 100 (47, [*M* - $COCH_2NCH_3$ + 2H]$^+$) and 84 (100, [*M* - CO_2 - CH_2NCH_3 + 3H]$^+$).

***N*-Benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methylamide (75)**



To a saturated solution of methylamine in dry methanol (30 cm^3) was added *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (**58**) (1.67 g, 5 mmol).

The solution was placed in a tightly stoppered vessel and left at room temperature. When the reaction was complete as judged by TLC (12 h), the stoppered vessel was cooled to 0 °C, opened and nitrogen gas bubbled through at room temperature for 5 min. The solvent was then removed under reduced pressure to give the product as a clear oil (1.62 g, 97%), *m/z* (Found: $[M + H]^+$, 334.1772. Calc. for $C_{17}H_{24}N_3O_4$: 334.1767); $[\alpha]_D^{22}$ -38.3 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3312 (NH), 1717 (carbamate CO), 1668 (tertiary amide CO), 1652 (methylamide CO) and 735 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.26 (*c*, 3H, d, *J* 8.0, CHC$\underline{H}_3$), 1.32 (*t*, 3H, d, *J* 8.0, CHC$\underline{H}_3$), 1.77-2.38 (*t* & *c*, 4H, m, γCH_2 and βCH_2), 2.72 (*t*, 3H, d, *J* 4.8, NHC$\underline{H}_3$), 2.78 (*c*, 3H, d, *J* 4.8, NHC$\underline{H}_3$), 3.41-3.74 (*t* & *c*, 2H, m, δCH_2), 4.26 (*t* &*c*, dd, J_1 4.5, J_2 6.2, αCH), 4.57 (*c*, 2H, d, *J* 8.2, CH_2Ph), 4.78 (*t*, 2H, d, *J* 6.4, CH_2Ph), 5.79 (*t* & *c*, 1H, d, *J* 8.0, N$\underline{H}$CHCH$_3$), 6.82 (*t* & *c*, 1H, m, N$\underline{H}$CH$_3$), 7.26 (*c*, 5H, s, aromatic) and 7.31 (*t*, 5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 17.45 (*c*, CHC$\underline{H}_3$), 18.59 (*t*, CHC$\underline{H}_3$), 22.47 (*c*, γCH_2), 25.48 (*t*, γCH_2), 26.63 (*t*, NHC$\underline{H}_3$), 26.96 (*c*, NHC$\underline{H}_3$), 28.48 (*t*, βCH_2), 32.07 (*c*, βCH_2), 47.40 (*t*, δCH_2), 47.71 (*c*, δCH_2), 48.83 (*t*, $\underline{C}HCH_3$), 49.50 (*c*, $\underline{C}HCH_3$), 60.34 (*t*, αCH), 61.31 (*c*, αCH), 67.18 (*t*, CH_2Ph), 67.31 (*c*, CH_2Ph), 128.47 (*t* & *c*, *para* aromatic), 128.58 (*t* & *c*, *ortho* aromatic), 128.94 (*t* & *c*, *meta* aromatic), 136.68 (*c*, quat. aromatic), 136.79 (*t*, quat. aromatic), 156.22 (*t*, $\underline{C}O_2CH_2Ph$), 156.89 (*c*, $\underline{C}O_2CH_2Ph$), 171.80 (*c*, $\underline{C}OCHCH_3$), 172.31 (*t*, $\underline{C}OCHCH_3$) and 172.78 (*t* & *c*, $\underline{C}ONHCH_3$); *m/z* (CI) 334 ($[M + H]^+$, 23%), 229 (8, $[M - CH_2Ph - CH_3 + 2H]^+$), 200 (3, $[M - CO_2CH_2Ph + 2H]^+$), 179 (13, $[PhCH_2CO_2NHCH_2CH_3]^+$) and 70 (100, $[C_4H_8N]^+$).

_N_-Benzyloxycarbonyl-(2*R*)-alanyl-(2*S*)-proline methylamide (76)

This compound was prepared in a manner identical with that for the methylamide (**75**), starting from *N*-benzyloxycarbonyl-(2*R*)-alanyl-(2*S*)-proline methyl ester (**59**) to give the product as a clear oil (1.56 g, 93%), *m/z* (Found: $[M + H]^+$, 334.1770. Calc. for $C_{17}H_{24}N_3O_4$: 334.1767); $[\alpha]_D^{32}$ -32.1 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3330 (NH), 1701 (carbamate CO), 1647 (tertiary amide CO), 1638 (methylamide CO) and 740 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.48 (3H, d, *J* 6.8, CHC$\underline{H}_3$), 2.14 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.48 (1H, m, $\frac{1}{2}CH_2$), 2.84 (3H, d, *J* 4.8, NHC$\underline{H}_3$), 3.59 (1H, m, $\frac{1}{2}\delta CH_2$), 3.99 (1H, m, $\frac{1}{2}\delta CH_2$), 4.56 (1H, m, C$\underline{H}$CH$_3$), 4.70 (1H, m, αCH), 5.24 (2H, s, CH_2Ph), 6.10 (1H, m, N$\underline{H}$CH$_3$), 7.20 (1H, d, *J* 2.4, N$\underline{H}$CHCH$_3$) and 7.48 (5H, m, aromatic); δ_C (74.76 MHz; C^2HCl_3) 16.25 (*t*, CHC$\underline{H}_3$), 17.51 (*c*, CHC$\underline{H}_3$), 22.22 (*c*, γCH_2), 24.02 (*t*, γCH_2), 25.83 (*t* & *c*, NCH_3), 28.87 (*t*, βCH_2), 31.83 (*c*, βCH_2), 46.76 (*t* & *c*, δCH_2), 48.04 (*c*, $\underline{C}HCH_3$), 48.61 (*t*, $\underline{C}HCH_3$), 60.46 (*t* & *c*,

αCH), 66.35 (*t* & *c*, CH_2Ph), 127.59 (*t* & *c*, *ortho* aromatic), 127.88 (*t* & *c*, *para* aromatic), 128.31 (*t* & *c*, *meta* aromatic), 136.35 (*t* & *c*, quat. aromatic), 155.98 (*c*, $\underline{C}O_2CH_2Ph$), 156.49 (*t*, $\underline{C}O_2CH_2Ph$), 171.24 (*c*, $\underline{C}OCHCH_3$), 171.87 (*t*, $\underline{C}OCHCH_3$), 172.10 (*t*, $\underline{C}ONHCH_3$) and 172.32 (*c*, $\underline{C}ONHCH_3$); *m/z* (CI) 334 ([*M* + H]$^+$, 15%), 244 (24, [*M* - CH_2Ph + 2H]$^+$), 212 (88, [*M* - OCH_2Ph - CH_3]$^+$, 200 (11, [*M* - CO_2CH_2Ph + 2H]$^+$), 97 (100, [C_4H_8NCO]$^+$) and 70 (13, [C_4H_8N]$^+$).

(2*S*)-Alanyl-(2*S*)-proline methylamide (77)

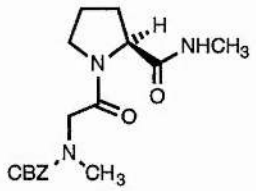
To a solution of *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methylamide (**75**) (0.67 g, 2 mmol) in methanol (30 cm^3) was added 10% palladium on activated charcoal (30 mg) and the vessel flushed with hydrogen gas. The resultant suspension was stirred vigorously under an atmosphere of hydrogen for 12 h. The mixture was then filtered through a prewashed celite pad and the solvent removed under reduced pressure to give a white solid which was recrystallised from methanol to give the product as a clear oil (0.39 g, 97%), *m/z* (Found: [*M* + H]$^+$, 200.1402. Calc. for $C_9H_{18}N_3O_2$: 200.1399); $[\alpha]_D^{22}$ +7.0 (*c* 0.1 in MeOH); υ_{max} (thin film)/cm^{-1} 3307 (NH), 1666 (tertiary amide CO) and 1652 (secondary amide CO); δ_H (200 MHz; $C^2H_3O^2H$) 1.21 (*c*, 3H, d, *J* 6.8, $CHC\underline{H}_3$), 1.28 (*t*, 3H, d, *J* 6.8 $CHC\underline{H}_3$), 1.82-2.40 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 2.74 (*t*, 3H, s, $NHC\underline{H}_3$), 2.78 (*c*, 3H, s, $NHC\underline{H}_3$), 3.54-3.81 (*t* & *c*, 3H, m, δCH_2 and $C\underline{H}CH_3$) and 4.37 (*t* & *c*, 1H, dd, J_1 4.6, J_2 8.0, αCH); δ_C (50.31 MHz; $C^2H_3O^2H$) 20.69 (*c*, $CH\underline{C}H_3$), 20.80 (*t*, $CH\underline{C}H_3$), 23.66 (*c*, γCH_2), 26.19 (*t*, γCH_2), 26.66 (*t*, $NHCH_3$), 26.91 (*c*, $NHCH_3$), 30.95 (*t*, βCH_2), 33.47 (*c*, βCH_2), 48.52 (*t*, δCH_2), 49.60 (*c*, δCH_2), 61.70 (*c*, αCH), 61.92 (*t*, αCH), 175.09 (*c*, $\underline{C}OCHCH_3$), 175.33 (*t*, $\underline{C}OCHCH_3$), 176.91 (*t*, $\underline{C}ONHCH_3$) and 177.47 (*c*, $\underline{C}ONHCH_3$); *m/z* (CI) 200 ([*M* + H]$^+$, 12%), 169 (23, [*M* - $NHCH_3$]$^+$), 156 (31, [*M* - $NHCH_3$ - CH_3 + 2H]$^+$), 142 (7, [*M* - $CONHCH_3$ + H]$^+$), 126 (13, [*M* - $CONHCH_3$ - CH_3]$^+$) and 70 (100, [C_4H_8N]$^+$).

(2*R*)-Alanyl-(2*S*)-proline methylamide (78)

This compound was prepared in a manner identical with that for the (2*S*, 2*S*) methylamide (**77**), using *N*-benzyloxycarbonyl-(2*R*)-alanyl-(2*S*)-proline methylamide (**76**) (0.67 g, 2 mmol) to give the product as a clear oil (0.41 g, 98%), *m/z* (Found: [*M* + H]$^+$, 200.1397. Calc. for $C_9H_{18}N_3O_2$: 200.1399); $[\alpha]_D^{22}$ -64.7 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3291 (NH), 1642 (tertiary amide CO) and 1638 (secondary

amide CO); δ_H (200 MHz; C^2HCl_3) 1.25 (3H, d, *J* 6.8, CHC<u>H</u>$_3$), 1.76-2.22 (3H, m, $\frac{1}{2}\beta CH_2$ and γCH_2), 2.22-2.55 (1H, m, $\frac{1}{2}\beta CH_2$), 2.74 (3H, d, *J* 4.8, NHC<u>H</u>$_3$), 3.42 (2H, m, δCH_2), 3.67 (1H, m, C<u>H</u>CH$_3$), 4.47 (1H, dd, J_1 2.4, J_2 7.8, αCH) and 7.14 (1H, m, N<u>H</u>CH$_3$); δ_C (50.31 MHz; $C^2H_3O^2H$) 20.11 (*c*, CH<u>C</u>H$_3$), 20.31 (*t*, CH<u>C</u>H$_3$), 23.75 (*c*, γCH_2), 25.80 (*t*, γCH_2), 26.75 (*t*, NHCH$_3$), 26.88 (*c*, NHCH$_3$), 31.10 (*t*, βCH_2), 33.39 (*c*, βCH_2), 48.41 (*t* & *c*, δCH_2), 49.72 (*t* & *c*, <u>C</u>HCH$_3$), 61.85 (*c*, αCH), 62.07 (*t*, αCH), 175.11 (*c*, <u>C</u>OCHCH$_3$), 175.19 (*t*, <u>C</u>OCHCH$_3$), 176.59 (*t*, <u>C</u>ONHCH$_3$) and 177.02 (*c*, <u>C</u>ONHCH$_3$); *m/z* (CI) 200 ([*M* + H]$^+$, 8%), 168 (18, [*M* - NH_2CH_3]$^+$), 140 (8, [*M* - $2NH_2CH_3$ + 3H]$^+$), 112 (10, [*M* - CONHCH$_3$ - NH_2CH_3 + 2H]$^+$), 70 (78, [C_4H_8N]$^+$) and 44 (100, [NH_2CH_3CH]$^+$).

_N_-Benzyloxycarbonyl-sarcosyl-(2*S*)-proline methylamide (79)



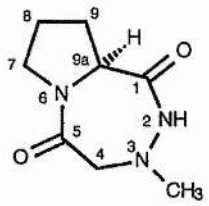
This compound was prepared in a manner identical with that for the *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methylamide (**75**), starting from *N*-benzyloxycarbonyl-sarcosyl-(2*S*)-proline methyl ester (**72**) (1.67 g, 5 mmol) to give the product as a clear oil (1.59 g, 96%), *m/z* (Found: [*M* + H]$^+$, 334.1761. Calc. for $C_{17}H_{24}N_3O_4$: 334.1766); $[\alpha]_D^{22}$ -58.0 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3328 (NH), 1701 (carbamate CO), 1672 (tertiary amide CO), 1663 (methylamide CO) and 733 & 699 (aromatic CH); δ_H (200 MHz; C^2HCl_3) 1.58-2.23 (*tt*, *ct* & *tc*, 3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.23-2.49 (*tt*, *ct* & *tc*, 1H, m, $\frac{1}{2}\beta CH_2$), 2.62 (*ct*, 3H, d, *J* 4.8, NHC<u>H</u>$_3$), 2.72 (*tt*, 3H, d, *J* 4.8, NHC<u>H</u>$_3$), 2.81 (*tc*, 3H, d, *J* 4.8, NHC<u>H</u>$_3$), 3.03 (*ct*, 3H, s, $CH_2NC\underline{H}_3$), 3.05 (*tt* & *tc*, 3H, s, $CH_2NC\underline{H}_3$), 3.15-3.68 (*tt*, *ct* & *tc*, 2H, m, δCH_2), 3.99 (*ct*, 2H, s, $COCH_2$), 4.00 (*tt*, 2H, s, $COCH_2$), 4.02 (*tc*, 2H, s, $COCH_2$), 4.29 (*tc*, 1H, dd, J_1 1.9, J_2 7.9, αCH), 4.53 (*tt* & *ct*, 1H, dd, $J_1 = J_2$ 6.0, αCH), 5.10 (*ct*, 2H, s, CH_2Ph), 5.13 (*tt* & *tc*, 2H, s, CH_2Ph), 6.91 (*tc*, 1H, m, N<u>H</u>CH$_3$), 7.01 (*tt* & *ct*, 1H, m, N<u>H</u>CH$_3$), 7.26 (*tc*, 5H, s, aromatic), 7.30 (*ct*, 5H, s, aromatic) and 7.34 (*tt*, 5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 22.29 (*tc*, γCH_2), 24.67 (*tt* & *ct*, γCH_2), 25.96 (*ct*, NHC<u>H</u>$_3$), 26.03 (*tt*, NHC<u>H</u>$_3$), 26.21 (*tc*, NHC<u>H</u>$_3$), 28.29 (*ct*, βCH_2), 28.63 (*tt*, βCH_2), 32.12 (*tc*, βCH_2), 35.72 (*tt* & *tc*, $CH_2NC\underline{H}_3$), 36.17 (*ct*, $CH_2NC\underline{H}_3$), 46.40 (*tt* & *ct*, δCH_2), 46.96 (*tc*, δCH_2), 50.80 (*ct*, $COCH_2$), 50.97 (*tc*, $COCH_2$), 51.30 (*tt*, $COCH_2$), 59.84 (*tc*, αCH), 60.10 (*ct*, αCH), 60.26 (*tt*, αCH), 66.99 (*ct* & *tc*, CH_2Ph), 67.13 (*tt*,

CH_2Ph), 127.37 (*ct & tc*, *para* aromatic), 127.46 (*tt*, *para* aromatic and *ct & tc*, *ortho* aromatic), 127.86 (*tt*, *ortho* aromatic), 128.36 (*tt*, *ct & tc*, *meta* aromatic), 136.45 (*tt*, quat. aromatic), 136.56 (*ct & tc*, quat. aromatic), 156.20 (*ct*, $\underline{C}OCH_2Ph$), 156.39 (*tc*, $\underline{C}OCH_2Ph$), 156.81 (*tt*, $\underline{C}OCH_2Ph$), 167.79 (*tc*, $\underline{C}OCH_2NCH_3$), 167.92 (*ct*, $\underline{C}OCH_2NCH_3$), 167.96 (*tt*, $\underline{C}OCH_2NCH_3$), 171.96 (*ct*, $\underline{C}ONHCH_3$), 172.06 (*tt*, $\underline{C}ONHCH_3$) and 172.16 (*tc*, $\underline{C}ONHCH_3$); *m/z* (CI) 334 ([*M* + H]$^+$, 100%), 229 (5, [*M* - OCH_2Ph + 3H]$^+$), 211 (3, [*M* - OCH_2Ph - CH_3]$^+$), 200 (8, [*M* - CO_2CH_2Ph + 2H]$^+$) and 91 (4, [CH_2Ph]$^+$).

Sarcosyl-(2*S*)-proline methylamide (80)

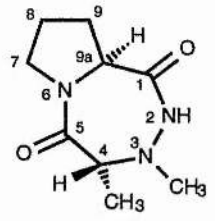
This compound was prepared in a manner identical with that for (2*S*)-alanyl-(2*S*)-proline methylamide (**77**), using *N*-benzyloxycarbonyl-sarcosyl-(2*S*)-proline methylamide (**79**) (0.67 g, 2 mmol) instead of *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-prolinamide (**75**) to give the product as a clear oil (0.38 g, 95%), *m/z* (Found: [*M* + H]$^+$, 200.1403. Calc. for $C_9H_{18}N_3O_2$: 200.1399); $[\alpha]_D^{22}$ -75.1 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3307 (NH), 1668 (tertiary amide CO) and 1652 (secondary amide CO); δ_H (200 MHz; C^2HCl_3) 1.62-2.55 (4H, m, βCH_2 and γCH_2), 2.47 (3H, s, $CONHC\underline{H}_3$), 2.54-3.05 (2H, m, δCH_2), 2.73 (3H, s, $CH_2NHC\underline{H}_3$), 2.91 (1H, m, $CH_2N\underline{H}CH_3$), 3.21-3.70 (2H, m, $COCH_2$), 4.54 (1H, dd, J_1 1.9, J_2 7.0, αCH) and 7.17 (1H, m, $CON\underline{H}CH_3$); δ_C (50.31 MHz; C^2HCl_3) 22.20 (*cc*, γCH_2), 22.49 (*tc*, γCH_2), 24.70 (*tt*, γCH_2), 25.01 (*ct*, γCH_2), 26.15 (*tt & tc*, $CONHC\underline{H}_3$), 26.30 (*ct & cc*, $CONHC\underline{H}_3$), 28.01 (*ct*, βCH_2), 28.63 (*tt*, βCH_2), 31.68 (*cc*, βCH_2), 32.17 (*tc*, βCH_2), 35.88 (*tc*, $CH_2NHC\underline{H}_3$), 36.05 (*tt*, $CH_2NHC\underline{H}_3$), 36.79 (*cc*, $CH_2NHC\underline{H}_3$), 37.29 (*ct*, $CH_2NHC\underline{H}_3$), 46.42 (*tt*, δCH_2), 46.84 (*cc*, δCH_2), 46.93 (*tc*, δCH_2), 47.03 (*ct*, δCH_2), 52.50 (*tc*, $CO\underline{C}H_2$), 52.83 (*tt*, $CO\underline{C}H_2$), 53.73 (*cc*, $CO\underline{C}H_2$), 53.86 (*ct*, $CO\underline{C}H_2$), 56.90 (*tc*, αCH), 57.73 (*cc*, αCH), 60.00 (*ct*, αCH), 60.13 (*tt*, αCH), 169.99 (*cc*, $\underline{C}OCH_2$), 170.29 (*tt & ct*, $\underline{C}OCH_2$), 170.79 (*tc*, $\underline{C}OCH_2$), 172.21 (*tt & ct*, $\underline{C}ONHCH_3$), 172.31 (*tc*, $\underline{C}ONHCH_3$) and 172.43 (*cc*, $\underline{C}ONHCH_3$); *m/z* (CI) 200 ([*M* + H]$^+$, 29%), 186 (40, [*M* - CH_3 + 2H]$^+$), 156 (62, [*M* - $NHCH_3$ - CH_3 + 2H]$^+$), 142 (6, [*M* - $CONHCH_3$ + H]$^+$), 128 (10, [*M* - $CONHCH_3$ - CH_3 + 2H]$^+$) and 70 (100, [C_4H_8N]$^+$).

(9a*S*)-3-Methyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (81)



To a stirred solution of methylhydrazine (0.46 g, 10 mmol) in CH_2Cl_2 (15 cm^3) was added *N*-bromoacetyl-(2*S*)-proline methyl ester (**38**) (1.25 g, 5 mmol). The mixture was refluxed for 30 min, allowed to cool to room temperature and extracted with distilled water (2 x 20 cm^3). The organic phase was separated, the solvent removed under reduced pressure and the residue redissolved in ethanol (15 cm^3). Methylhydrazine (0.23 g, 5 mmol) was added and the resultant solution refluxed for 1 h. Removal of the solvent under reduced pressure gave a white solid which was recrystallised from methanol to give the product as colourless crystals (0.75 g, 82%), m.p. 236-8 °C; (Found: C, 52.3; H, 7.3; N, 22.85. Calc. for $C_8H_{13}N_3O_2$: C, 52.45; H, 7.2; N, 22.95%); $[\alpha]_D^{22}$ +118.1 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3101 (NH), 1699 (tertiary amide CO) and 1597 (secondary amide CO); δ_H (200 MHz; $C^2H_3O^2H$), 1.76-2.15 (3H, m, $\frac{1}{2}\beta CH_2$ and γCH_2), 2.33-2.58 (1H, m, $\frac{1}{2}\beta CH_2$), 2.69 (3H, s, NCH_3), 3.41 (1H, d, *J* 17.2, $\frac{1}{2}COCH_2$), 3.55 (2H, dd, $J_1 = J_2$ 6.6, δCH_2), 3.71 (1H, d, *J* 17.2, $\frac{1}{2}COCH_2$) and 5.19 (1H, dd, $J_1 = J_2$ 7.1, αCH); δ_C (50.31 MHz; $C^2H_3O^2H$) 22.63 (γCH_2), 28.05 (βCH_2), 44.12 (NCH_3), 48.46 (δCH_2), 58.50 ($CO\underline{C}H_2$), 63.55 (αCH), 169.77 ($\underline{C}OCH_2$) and 173.36 (CONH); *m/z* (FAB) 184 ([*M* + H]$^+$, 100%), 168 (7, [*M* - CH_3]$^+$), 154 (19, [*M* - NCH_3]$^+$), 112 (26, [*M* - $CONHNCH_3$ + H]$^+$) and 70 (63, [C_4H_8N]$^+$).

(4*R*, 9a*S*)-3,4-Dimethyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (82)



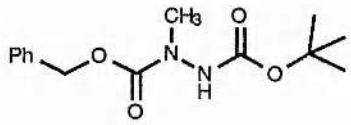
This compound was prepared in a manner identical with that for the triazepine-1,5-dione (**81**), using *N*-(2*S*)-chloropropanoyl-(2*S*)-proline methyl ester (**48**) (1.10 g, 5 mmol) instead of *N*-bromoacetyl-(2*S*)-proline methyl ester (**38**) to give the product as colourless crystals (0.83 g, 84%), m.p. 155-8 °C; (Found: C, 54.6; H, 7.85; N, 21.4. Calc. for $C_9H_{15}N_3O_2$: C, 54.8; H, 7.7; N, 21.3%); *m/z* (Found: $[M + H]^+$, 198.1244. Calc. for $C_9H_{16}N_3O_2$: 198.1243); $[\alpha]_D^{22}$ +104.2 (*c* 0.1 in MeOH); υ_{max} (Nujol)/cm^{-1} 3184 & 3124 (NH), 1695 (tertiary amide CO) and 1600 (secondary amide CO); δ_H (300 MHz; C^2HCl_3) 1.38 (3H, d, *J* 6.6, CHC$\underline{H}_3$), 1.67-2.02 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.39-2.75 (1H, m, $\frac{1}{2}\beta CH_2$), 2.58 (3H, s, NCH_3), 3.37 (1H, q, *J* 6.6, C$\underline{H}$CH$_3$), 3.35-3.66 (1H, m, $\frac{1}{2}\delta CH_2$), 3.80 (1H, m, $\frac{1}{2}\delta CH_2$), 5.10 (1H, dd, J_1 5.5, J_2 7.6, αCH) and 7.69 (1H, s, NH); δ_C (74.76 MHz; C^2HCl_3) 18.69 (CH$\underline{C}$H$_3$), 22.91 (γCH_2), 27.01 (βCH_2), 41.80 (NCH_3), 48.76 (δCH_2), 57.04 (αCH), 68.64 ($\underline{C}$HCH$_3$), 171.69 ($\underline{C}$OCHCH$_3$) and 173.09 (CONH); *m/z* (FAB) 198 ($[M + H]^+$, 100%), 182 (6, $[M - CH_3]^+$), 168 (24, $[M - NCH_3]^+$), 126 (19, $[M - CONHCH_3 + H]^+$) and 70 (85, $[C_4H_8N]^+$).

(4*S*, 9a*S*)-3,4-Dimethyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (83)

This compound was prepared in a manner identical with that for the (4*R*, 9a*S*) diastereomer (**82**), using *N*-(2*R*)-chloropropanoyl-(2*S*)-proline methyl ester (**49**) (1.10 g, 5 mmol) to give the product as colourless crystals (0.79 g, 80%), m.p. 146-50 °C; (Found: C, 54.6; H, 7.9; N, 21.4. Calc. for $C_9H_{15}N_3O_2$: C, 54.8; H, 7.7; N, 21.3%); *m/z* (Found: $[M + H]^+$, 198.1244. Calc. for $C_9H_{16}N_3O_2$: 198.1243); $[\alpha]_D^{22}$ +18.2 (*c* 0.1 in MeOH); υ_{max} (Nujol)/cm^{-1} 2926 & 2854 (NH), 1696 (tertiary amide CO) and 1601 (secondary amide CO); δ_H (200 MHz; C^2HCl_3) 1.31 (3H, d, *J* 7.0, CHC$\underline{H}_3$), 1.71-2.12 (3H, m, $\frac{1}{2}\beta CH_2$ and γCH_2), 2.38-2.62 (1H, m, $\frac{1}{2}\beta CH_2$), 2.67 (3H, s, NCH_3), 3.60 (2H, m, δCH_2), 3.65 (1H, q, *J* 7.0, C$\underline{H}$CH$_3$), 5.00 (1H, dd,

$J_1 = J_2$ 7.6, αCH) and 7.24 (1H, s, NH); δ_C (50.31 MHz; C^2HCl_3) 9.21 ($CH\underline{C}H_3$), 22.70 (γCH_2), 27.69 (βCH_2), 41.88 (NCH_3), 48.57 (δCH_2), 58.10 (αCH), 65.08 ($\underline{C}HCH_3$), 170.80 ($\underline{C}OCHCH_3$) and 172.65 (CONH); *m/z* (FAB) 198 ([*M* + H]$^+$, 100%), 182 (6, [*M* - CH_3]$^+$), 168 (24, [*M* - NCH_3]$^+$), 126 (19, [*M* - $CONHCH_3$ + H]$^+$) and 70 (85, [C_4H_8N]$^+$).

_N_-*tert*-Butoxycarbonyl-_N'_-benzyloxycarbonyl-_N'_-methylhydrazine (86)



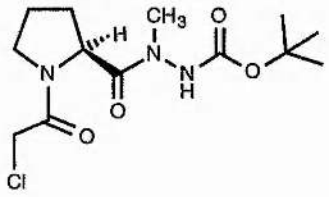
To a solution of *N*-benzyloxycarbonyl-*N*-methylhydrazine (3.60 g, 20 mmol) in *iso*propanol (25 cm^3) was added a solution of ditertiarybutyldicarbonate (4.80 g, 22 mmol) in CH_2Cl_2 (10 cm^3). The mixture was stirred for 4 h, and the solvent removed under reduced pressure to yield the product as a pale yellow oil. Purification by silica column chromatography using ethyl acetate/ light petroleum (1:4) as the eluent gave the pure product as a clear oil (4.94 g, 88%), *m/z* (Found: [*M* + H]$^+$, 281.1504. Calc. for $C_{14}H_{21}N_2O_4$: 281.1501); υ_{max} (thin film)/cm^{-1} 3307 (NH), 1733 (CBZ CO) and 1707 (^{t}BOC CO); δ_H (200 MHz; C^2HCl_3) 1.29 (9H, s, $C(CH_3)_3$), 2.45 (3H, s, NCH_3), 5.12 (2H, s, CH_2Ph), 6.12 (1H, s, NH) and 7.31 (5H, s, aromatic); δ_C (50.31 MHz; C^2HCl_3) 28.31 ($C(\underline{C}H_3)_3$), 39.25 ($NHCH_3$), 63.98 (CH_2Ph), 80.74 ($\underline{C}(CH_3)_3$), 128.15 (*ortho* aromatic), 128.33 (*meta* aromatic), 128.78 (*para* aromatic), 136.43 (quat. aromatic), 155.52 ($\underline{C}O_2CH_2Ph$) and 157.11 ($CO_2{}^tBu$); *m/z* (CI) 281 ([*M* + H]$^+$, 54%), 265 (32, [*M* - CH_3]$^+$), 225 (71, [*M* - $C(CH_3)_3$ + 2H]$^+$), 191 (57, [*M* - $CO_2C(CH_3)_3$ + 2H]$^+$), 91 (100, [$PhCH_2$]$^+$) and 57 (88, [$C(CH_3)_3$]$^+$).

_N_-*tert*-Butoxycarbonyl-_N'_-methylhydrazine (87)

To a solution of *N*-*tert*-butoxycarbonyl-*N'*-benzyloxycarbonyl-*N'*-methylhydrazine (**86**) (1.40 g, 5 mmol) in methanol (30 cm^3) was added 10% palladium on activated charcoal (75 mg) and the vessel flushed with hydrogen gas. The resultant suspension was stirred vigorously under an atmosphere of hydrogen for 12 h. The mixture was then filtered through a prewashed celite pad, and the solvent removed under reduced pressure to give a clear oil which solidified on standing to give colourless crystals of the product (0.68 g, 93%), m.p. 46-9 °C; (Found: C, 49.25; H, 9.9; N, 19.2. Calc. for $C_6H_{14}N_2O_2$: C, 49.3 H, 9.65; N, 19.15%); υ_{max} (Nujol)/cm^{-1} 3320 (amine NH), 3241

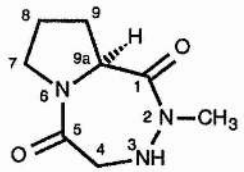
(amide NH) and 1700 (CO); δ_H (200 MHz; C^2HCl_3) 1.33 (9H, s, $C(CH_3)_3$), 2.48 (3H, s, $NHC\underline{H}_3$), 3.76 (1H, s, $N\underline{H}CH_3$) and 6.00 (1H, s, CONH); δ_C (50.31 MHz; C^2HCl_3) 28.88 ($C(\underline{C}H_3)_3$), 39.69 ($NHCH_3$), 80.82 ($\underline{C}(CH_3)_3$) and 157.27 ($CO_2{}^tBu$); *m/z* (EI) 146 (M^+, 15%), 131 (3, $[M - CH_3]^+$), 103 (51, $[M - NHNHCH_3 + 2H]^+$), 90 (33, $[M - C(CH_3)_3 + H]^+$), 73 (27, $[OC(CH_3)_3]^+$) and 57 (100, $[C(CH_3)_3]^+$).

N-tert-Butoxycarbonyl-_N'_-(_N_-chloroacetyl-(2_S_)-prolyl)-_N'_-methylhydrazine (89)



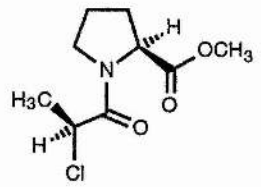
To a solution of *N*-chloroacetyl-(2*S*)-proline (**88**) (1.92 g, 10 mmol) in dry THF (25 cm^3) was added *N*-methylmorpholine (1.12 cm^3, 10 mmol) and the solution cooled to -15 °C. *iso*-Butyl chloroformate (1.36 cm^3, 10 mmol) was added with stirring and the solution left stirred at -15 °C for 2 min. A solution of *N-tert*-butoxycarbonyl-*N'*-methylhydrazine (**87**) (1.46 g, 10 mmol) in dry THF (20 cm^3) was then added. The reaction mixture was allowed to warm up to room temperature and left to stir for 6 h. The hydrochloride salts were filtered off and the solvents removed under reduced pressure. The resultant oily residue was dissolved in CH_2Cl_2 (25 cm^3) and washed with 0.5 mol dm^{-3} HCl (2 x 15 cm^3) and 5% sodium carbonate solution (2 x 15 cm^3). The organic phase was then dried ($MgSO_4$) and the solvent removed under reduced pressure to give a yellow oil. The crude material was purified by silica column chromatography using ethyl acetate as the eluent to give the product as a white solid (2.17 g, 68%), m.p. 64-7 °C; *m/z* (Found: $[M + H]^+$, 320.1373. Calc. for $C_{13}H_{23}{}^{35}ClN_3O_4$: 320.1376); $[\alpha]_D^{22}$ -11.0 (*c* 0.5 in MeOH); υ_{max} (Nujol)/cm^{-1} 3227 (NH), 1730 (carbamate CO), 1695 (methylamide CO), 1678 (tertiary amide CO) and 771 (C-Cl); δ_H (200 MHz; C^2HCl_3) 1.48 (9H, s, $C(CH_3)_3$), 1.74-2.45 (4H, m, γCH_2 and βCH_2), 3.11 (3H, s, NCH_3), 3.63 (2H, m, δCH_2), 4.04 (1H, d, *J* 13.0, $\frac{1}{2}COCH_2$), 4.13 (1H, d, *J* 13.0, $\frac{1}{2}COCH_2$), 4.97 (1H, dd, J_1 5.6, J_2 6.2, αCH) and 7.90 (1H, s, NH); δ_C (50.31 MHz; C^2HCl_3) 25.15 (γCH_2), 28.47 ($C(\underline{C}H_3)_3$), 28.55 (βCH_2), 35.54 (NCH_3), 42.48 ($CO\underline{C}H_2$), 47.59 (δCH_2), 57.32 (αCH), 81.49 ($\underline{C}(CH_3)_3$), 154.96 (CONH), 165.31 ($\underline{C}ONCH_3$) and 173.73 ($\underline{C}OCH_2$); *m/z* (CI) 320 ($[M + H]^+$, 3%), 284 (100, $[M - Cl]^+$), 240 (91, $[M - Cl - C(CH_3)_3 + H]^+$), 225 (9, $[M - Cl - OC(CH_3)_3 + 2H]^+$), 170 (15, $[M - CH_2Cl - CO_2C(CH_3)_3 + H]^+$) and 155 (4, $[M - CH_2Cl - NHCO_2C(CH_3)_3 + H]^+$).

(9a*S*)-2-Methyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (91)



Hydrogen chloride gas was bubbled through a solution of *N-tert*-butoxycarbonyl-*N'*-(*N*-chloroacetyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**89**) (1.60 g, 5 mmol) in ethylacetate (30 cm^3) for 20 min at 0 °C. The solvent was then removed under reduced pressure and the resultant hygroscopic white solid dissolved in water (25 cm^3) and NaOH solution (1 mol dm^{-3}) added until the solution reached pH 9. The solution was then extracted with ethyl acetate (3 x 20 cm^3), the combined organic fractions were dried ($MgSO_4$) and the solvent removed under reduced pressure to give the product as a white solid (0.59 g, 64%), m.p. 75-9 °C; *m/z* (Found: M^+, 183.1001. Calc. for $C_8H_{13}N_3O_2$: 183.1007); $[\alpha]_D^{22}$ -40.2 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3440 (NH), 1652 (methylamide CO) and 1637 (tertiary amide CO); δ_H (200 MHz; C^2HCl_3) 1.82-2.38 (*A* & *B*, 4H, m, γCH_2 and βCH_2), 3.18 (*A*, 3H, s, NCH_3), 3.22 (*B*, 3H, s, NCH_3), 3.69 (*A* & *B*, 2H, m, δCH_2), 3.91 (*A* & *B*, 1H, s, NH), 4.03 (*A* & *B*, 1H, d, *J* 12.4, $\frac{1}{2}COCH_2$), 4.13 (*A* & *B*, 1H, d, *J* 12.4, $\frac{1}{2}COCH_2$), 5.39 (*B*, 1H, dd, J_1 3.2, J_2 8.6, αCH) and 5.47 (*A*, 1H, dd, J_1 3.9, J_2 7.9, αCH); δ_C (50.31 MHz; C^2HCl_3) 22.22 (*B*, γCH_2), 24.52 (*A*, γCH_2), 29.07 (*A*, βCH_2), 31.67 (*B*, βCH_2), 38.26 (*A*, NCH_3), 38.43 (*B*, NCH_3), 42.41 (*A* & *B*, $CO\underline{C}H_2$), 47.44 (*A*, δCH_2), 47.65 (*B*, δCH_2), 57.46 (*A*, αCH), 57.88 (*B*, αCH), 164.51 (*A*, $\underline{C}ONCH_3$), 165.13 (*B*, $\underline{C}ONCH_3$) and 172.97 (*A* & *B*, $\underline{C}OCH_2$); *m/z* (EI) 183 (M^+, 45%), 155 (15, $[M - NCH_3 + H]^+$), 139 (65, $[M - NHNCH_3]^+$), 125 (20, $[M - CH_2NHNCH_3]^+$), 111 (40, $[M - CONCH_3NH]^+$) and 70 (100, $[C_4H_8N]^+$).

_N_-(2_S_)-Chloropropanoyl-(2_S_)-proline methyl ester (94)



To a solution of _N_-methylmorpholine (1.12 cm^3, 10 mmol) in dry THF (20 cm^3) was added (2_S_)-chloropropionic acid (**92**),[228] (1.09 g, 10 mmol) and the solution cooled to -15 °C. *iso*-Butyl chloroformate (1.36 cm^3, 10 mmol) was added with stirring and the resultant suspension stirred at -15 °C for 2 min. A mixture of (2_S_)-proline methyl ester hydrochloride (**21**) (1.66 g, 10 mmol) and _N_-methylmorpholine (1.12 cm^3, 10 mmol) in dry DMF (5 cm^3) was added. The reaction mixture was allowed to warm up to room temperature and left to stir for 3 h. The hydrochloride salts were filtered off and the solvents removed under reduced pressure. The resultant clear oil was dissolved in CH_2Cl_2 (25 cm^3) and washed with 0.5 mol dm^{-3} HCl (2 x 15 cm^3) and 5% sodium carbonate solution (2 x 15 cm^3). The organic phase was then dried ($MgSO_4$) and the solvent removed under reduced pressure to yield a yellow oil which was purified by silica chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a clear oil (1.56 g, 71%), *m/z* (Found: $[M + H]^+$, 220.0747. Calc. for $C_9H_{15}{}^{35}ClNO_3$: 220.0742); $[\alpha]_D^{22}$ -70.8 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 1746 (ester CO), 1659 (amide CO) and 1433 (C-O); δ_H (200 MHz; C^2HCl_3) 1.62 (3H, d, *J* 6.8, $CHC\underline{H}_3$), 1.85-2.32 (4H, m, γCH_2 and βCH_2), 3.51-3.92 (2H, m, δCH_2), 3.69 (3H, s, OCH_3), 4.47 (1H, dd, J_1 3.4, J_2 6.6, αCH) and 4.49 (1H, q, *J* 6.8, $C\underline{H}CH_3$); δ_C (50.31 MHz; C^2HCl_3) 20.50 (*t*, $CH\underline{C}H_3$), 21.32 (*c*, $CH\underline{C}H_3$), 21.99 (*c*, γCH_2), 24.79 (*t*, γCH_2), 29.05 (*t*, βCH_2), 31.35 (*c*, βCH_2), 46.91 (*t* & *c*, OCH_3), 51.17 (*t* & *c*, δCH_2), 52.13 (*t*, $\underline{C}HCH_3$), 52.48 (*c*, $\underline{C}HCH_3$), 59.18 (*t*, αCH), 59.28 (*c*, αCH), 167.52 (*t* & *c*, $\underline{C}OCHCH_3$), 172.07 (*c*, $\underline{C}O_2CH_3$) and 172.23 (*t*, $\underline{C}O_2CH_3$); *m/z* (EI) 220 ($[M + H]^+$, 7%), 184 (5, $[M - Cl]^+$), 170 (54, $[M - Cl - CH_3 + H]^+$), 160 (84, $[M - CO_2CH_3]^+$), 128 (58, $[M - Cl - CO_2CH_3 + 3H]^+$) and 70 (100, $[C_4H_8N]^+$).

_N_-(2_R_)-Chloropropanoyl-(2_S_)-proline methyl ester (95)

This compound was prepared in a manner identical with that for the (2_S_, 2_S_) diastereomer (**94**), using (2_R_)-chloropropionic acid (**93**),[228] (1.09 g, 10 mmol) to give the product as colourless crystals (1.82 g, 83%), m.p. 118-20 °C; (Found: C, 49.2; H,

6.4; N, 6.4. Calc. for $C_9H_{14}ClNO_3$: C, 49.05; H, 6.6; N, 6.4%); $[\alpha]_D^{22}$ -113.9 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 1751 (ester CO), 1656 (amide CO) and 1454 (C-O); δ_H (200 MHz; C^2HCl_3) 1.64 (*c*, 3H, d, *J* 6.4, $CHC\underline{H}_3$), 1.66 (*t*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.80-2.33 (*t* & *c*, 4H, m, γCH_2 and βCH_2), 3.46-3.70 (*t*, 2H, m, δCH_2), 3.74 (*t*, 3H, s, OCH_3), 3.77 (*c*, 3H, s, OCH_3), 3.82-3.98 (*c*, 2H, m, δCH_2), 4.24 (*c*, 1H, q, *J* 6.4, $C\underline{H}CH_3$), 4.46 (*t*, 1H, dd, J_1 3.2, J_2 8.4, αCH) and 4.71 (*c*, 1H, dd, J_1 5.0, J_2 5.6, αCH); δ_C (74.76 MHz; C^2HCl_3) 20.82 (*t*, $CH\underline{C}H_3$), 20.94 (*c*, $CH\underline{C}H_3$), 22.78 (*c*, γCH_2), 25.24 (*t*, γCH_2), 29.41 (*t*, βCH_2), 31.44 (*c*, βCH_2), 47.26 (*c*, δCH_2), 47.35 (*t*, δCH_2), 51.29 (*t*, OCH_3), 51.39 (*c*, OCH_3), 52.62 (*t*, $\underline{C}HCH_3$), 53.27 (*c*, $\underline{C}HCH_3$), 59.47 (*c*, αCH), 59.62 (*t*, αCH), 167.99 (*t*, $\underline{C}OCHCH_3$), 168.35 (*c*, $\underline{C}OCHCH_3$), 172.48 (*t*, $\underline{C}O_2CH_3$) and 172.78 (*c*, $\underline{C}O_2CH_3$); *m/z* (EI) 220 ($[M + H]^+$, 3%), 184 (4, $[M - Cl]^+$), 170 (5, $[M - Cl - CH_3 + H]^+$), 160 (63, $[M - CO_2CH_3]^+$), 127 (58, $[M - Cl - CO_2CH_3 + 2H]^+$) and 70 (100, $[C_4H_8N]^+$).

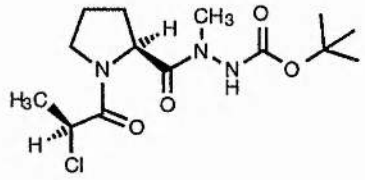
N-(2*S*)-Chloropropanoyl-(2*S*)-proline (96)

To a solution of *N*-(2*S*)-chloropropanoyl-(2*S*)-proline methyl ester (**94**) (2.20 g, 10 mmol) in methanol (20 cm^3) was added 1 mol dm^{-3} sodium hydroxide solution (22 cm^3). The resultant solution was stirred at room temperature for 1 h and then 1 mol dm^{-3} HCl solution (10 cm^3) added. Methanol was removed under reduced pressure and a second portion of 1 mol dm^{-3} HCl solution (10 cm^3) was added. The precipitated white solid was filtered and recrystallized from methanol to give the product as colourless crystals (1.77 g, 84%), m.p. 164-5 °C; (Found: C, 46.95; H, 5.95; N, 6.75. Calc. for $C_8H_{12}ClNO_3$: C, 46.75; H, 5.9; N, 6.8%); $[\alpha]_D^{22}$ -62.1 (*c* 1.0 in MeOH); υ_{max} (Nujol mull)/cm^{-1} 3045 (OH), 1739 (acid CO) and 1624 (amide CO); δ_H (200 MHz; C^2HCl_3) 1.68 (3H, d, *J* 6.6, CH_3), 1.92-2.33 (4H, m, βCH_2 and γCH_2), 3.64 (1H, m, $\frac{1}{2}\delta CH_2$), 3.78 (1H, m, $\frac{1}{2}\delta CH_2$), 4.53 (1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.57 (1H, dd, $J_1 = J_2$ 5.8, αCH) and 9.19 (1H, s, CO_2H); δ_C (50.31 MHz; C^2HCl_3) 20.99 (CH_3), 25.34 (γCH_2), 29.39 (βCH_2), 47.79 (δCH_2), 51.67 ($\underline{C}HCH_3$), 60.11 (αCH), 169.23 (NCO) and 175.57 (CO_2H); *m/z* (EI) 205 (M^+, 14%), 172 (46, $[M - Cl + 2H]^+$), 144 (66, $[M - CHClCH_3 + 2H]^+$), 116 (59, $[M - COCHClCH_3 + 2H]^+$), 97 (87, $[C_4H_8NCO]^+$) and 70 (100, $[C_4H_8N]^+$).

_N_-(2_R_)-Chloropropanoyl-(2_S_)-proline (97)

To a solution of _N_-(2_R_)-chloropropanoyl-(2_S_)-proline methyl ester (**95**) (2.20 g, 10 mmol) in methanol (20 cm^3) was added 1 mol dm^{-3} sodium hydroxide solution (22 cm^3). The resultant solution was stirred at room temperature for 1 h and then 1 mol dm^{-3} HCl solution (10 cm^3) added. Methanol was removed under reduced pressure, a second portion of 1 mol dm^{-3} HCl solution (10 cm^3) was added and the solution was extracted with ethyl acetate (3 x 50 cm^3). The combined organic fractions were dried ($MgSO_4$) and the solvent removed under reduced pressure to give the product as a white solid (1.54 g, 79%), m.p. 102-6 °C; _m/z_ (Found: $[M + H]^+$, 206.0589. Calc. for $C_8H_{13}{}^{35}ClNO_3$: 206.0584); $[\alpha]_D^{22}$ -110.2 (_c_ 1.0 in MeOH); υ_{max} (Nujol mull)/cm^{-1} 3344 (OH), 1750 (acid CO) and 1616 (amide CO); δ_H (200 MHz; C^2HCl_3) 1.64 (_c_, 3H, d, _J_ 6.4, CH_3), 1.65 (_t_, 3H, d, _J_ 6.6, CH_3), 1.85-2.43 (_t_ & _c_, 4H, m, βCH_2 and γCH_2), 3.58 (_t_ & _c_, 1H, m, $\frac{1}{2}\delta CH_2$), 3.89 (_t_ & _c_, 1H, m, $\frac{1}{2}\delta CH_2$), 4.32 (_c_, 1H, q, _J_ 6.4, C<u>H</u>CH_3), 4.48 (_t_, 1H, q, _J_ 6.6, C<u>H</u>CH_3), 4.52 (_t_, 1H, dd, J_1 4.3, J_2 6.7, αCH), 4.72 (_c_, 1H, dd, J_1 4.3, J_2 6.7, αCH), and 10.25 (_t_ & _c_, 1H, s, CO_2H); δ_C (50.31 MHz; C^2HCl_3) 20.45 (_t_ & _c_, CH_3), 22.37 (_c_, γCH_2), 24.70 (_t_, γCH_2), 28.88 (_t_, βCH_2), 30.99 (_c_, βCH_2), 47.21 (_c_, δCH_2), 47.26 (_t_, δCH_2), 51.07 (_t_ & _c_, <u>C</u>HCH_3), 59.36 (_t_ & _c_, αCH), 168.53 (_t_, NCO), 168.81 (_c_, NCO), 173.92 (_c_, CO_2H) and 174.73 (_t_, CO_2H); _m/z_ (CI) 206 ($[M + H]^+$, 58%), 170 (100, $[M - Cl]^+$), 156 (23, $[M - Cl - CH_3 + H]^+$), 142 (6, $[M - Cl - CHCH_3]^+$), 113 (12, $[M - CHCl - CO_2]^+$) and 97 (3, $[C_4H_8NCO]^+$).

N-tert-Butoxycarbonyl-_N′_-(_N_-(2_S_)-chloropropanoyl-(2_S_)-prolyl)-_N′_-methylhydrazine (98)



To a solution of _N_-(2_S_)-chloropropanoyl-(2_S_)-proline (**96**) (2.06 g, 10 mmol) in dry THF (40 cm^3) was added _N_-methylmorpholine (1.12 cm^3, 10 mmol) and the solution cooled to -15 °C. _iso_-Butyl chloroformate (1.36 cm^3, 10 mmol) was added with stirring and the resultant suspension stirred at -15 °C for 2 min. A solution of _N-tert_-butoxycarbonyl-_N′_-methylhydrazine (**87**) (1.46 g, 10 mmol) in dry THF (10 cm^3) was added and the reaction mixture was allowed to warm up to room temperature and left

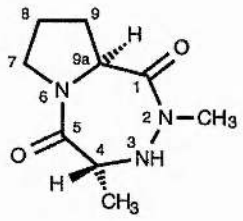
to stir for 12 h. The hydrochloride salts were filtered off and the solvent removed under reduced pressure. The resultant oily residue was dissolved in CH_2Cl_2 (25 cm^3) and washed with 0.5 mol dm^{-3} HCl (2 x 15 cm^3) and 5% sodium carbonate solution (2 x 15 cm^3). The organic phase was then dried ($MgSO_4$) and the solvent removed under reduced pressure to yield a yellow oil which was purified by silica chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a white solid (2.31 g, 69%), m.p. 81-3 °C; (Found: C, 50.6; H, 7.25; N, 12.5. Calc. for $C_{14}H_{24}ClN_3O_4$: C, 50.4; H, 7.25; N, 12.6%); *m/z* (Found: $[M + H]^+$, 334.1538. Calc. for $C_{14}H_{25}{}^{35}ClN_3O_4$: 334.1536); $[\alpha]_D^{22}$ -16.7 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3327 (NH), 1735 (carbamate CO), 1685 (methylamide CO) and 1646 (tertiary amide CO); δ_H (200 MHz; C^2HCl_3) 1.48 (9H, s, $C(CH_3)_3$), 1.64 (3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.82-2.31 (4H, m, γCH_2 and βCH_2), 3.12 (3H, s, NCH_3), 3.52-3.89 (2H, m, δCH_2), 4.53 (1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.98 (1H, dd, $J_1 = J_2$ 6.3, αCH) and 7.87 (1H, s, NH); δ_C (50.31 MHz; C^2HCl_3) 21.10 ($CH\underline{C}H_3$), 25.20 (γCH_2), 28.64 ($C(\underline{C}H_3)_3$), 28.78 (βCH_2), 35.55 (NCH_3), 47.77 (δCH_2), 51.68 ($\underline{C}HCH_3$), 57.77 (αCH), 81.78 ($\underline{C}(CH_3)_3$), 155.11 ($CO_2{}^tBu$), 168.09 ($\underline{C}ONCH_3$) and 173.86 ($\underline{C}OCHCH_3$); *m/z* (CI) 334 ($[M + H]^+$, 33%), 300 (15, $[M - Cl + 2H]^+$), 278 (100, $[M - C(CH_3)_3 + 2H]^+$), 244 (28, $[M - Cl - C(CH_3)_3 + 3H]^+$), 234 (12, $[M - CO_2C(CH_3)_3 + 2H]^+$) and 200 (13, $[M - Cl - CO_2C(CH_3)_3 + 3H]^+$).

N-tert-Butoxycarbonyl-_N'_-(_N_-(2_R_)-chloropropanoyl-(2_S_)-prolyl)-_N'_-methylhydrazine (99)

This compound was prepared in a manner identical with that for the (2*S*, 2*S*) diastereomer (**98**), using *N*-(2*S*)-chloropropanoyl-(2*S*)-proline (**87**) (2.06 g, 10 mmol) to give the product as a white solid (2.47 g, 74%), m.p. 132-5 °C; (Found: C, 50.6; H, 7.3; N, 12.5. Calc. for $C_{14}H_{24}ClN_3O_4$: C, 50.4; H, 7.25; N, 12.6%); *m/z* (Found: $[M + H]^+$, 334.1539. Calc. for $C_{14}H_{25}{}^{35}ClN_3O_4$: 334.1536); $[\alpha]_D^{22}$ -20.3 (*c* 0.2 in MeOH); υ_{max} (Nujol)/cm^{-1} 3209 (NH), 1728 (carbamate CO), 1687 (methylamide CO) and 1652 (tertiary amide CO); δ_H (200 MHz; C^2HCl_3) 1.47 (*t*, 9H, s, $C(CH_3)_3$), 1.51 (*c*, 9H, s, $C(CH_3)_3$), 1.60 (*c*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.65-2.48 (*t* & *c*, 4H, m, γCH_2 and βCH_2), 1.65 (*t*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 3.11 (*t*, 3H, s, NCH_3), 3.13 (*c*, 3H, s, NCH_3), 3.57 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 3.89 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 4.49 (*t* & *c*, 1H, q, *J* 6.8, $C\underline{H}CH_3$), 4.96 (*t* & *c*, 1H, dd, J_1 4.2, J_2 7.8, αCH) and 7.92 (*t* & *c*, 1H, s, NH); δ_C (50.31 MHz; C^2HCl_3) 20.75 (*c*, $CH\underline{C}H_3$), 21.05 (*t*, $CH\underline{C}H_3$), 22.37 (*c*, γCH_2), 24.71 (*t*, γCH_2), 28.03 (*c*, $C(\underline{C}H_3)_3$), 28.39 (*t*, $C(\underline{C}H_3)_3$), 31.37 (*c*, βCH_2), 35.17 (*t*, NCH_3), 36.05 (*c*, NCH_3), 47.37 (*c*, δCH_2), 47.50 (*t*, δCH_2), 51.15 (*t*, $\underline{C}HCH_3$),

53.86 (*c*, $\underline{C}HCH_3$), 57.25 (*t*, αCH), 58.02 (*c*, αCH), 81.38 (*t* & *c*, $\underline{C}(CH_3)_3$), 155.17 (*t* & *c*, $CO_2{}^tBu$), 167.95 (*t* & *c*, $\underline{C}ONCH_3$) and 173.86 (*t* & *c*, $\underline{C}OCHCH_3$); *m/z* (CI) 334 ([*M* + H]$^+$, 41%), 300 (11, [*M* - Cl + 2H]$^+$), 278 (100, [*M* - $C(CH_3)_3$ + 2H]$^+$), 234 (12, [*M* - $CO_2C(CH_3)_3$ + 2H]$^+$), 219 (11, [*M* - $NHCO_2C(CH_3)_3$ + 2H]$^+$) and 185 (8, [*M* - Cl - CH_3 - $CO_2C(CH_3)_3$ + 3H]$^+$).

(4*R*, 9a*S*)-2,4-Dimethyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (102)



Hydrogen chloride gas was bubbled through a solution of *N-tert*-butoxycarbonyl-*N'*-(*N*-(2*S*)-chloropropanoyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**98**) (1.35 g, 5 mmol) in ethylacetate (30 cm^3) for 20 min at 0 °C. The solvent was then removed under reduced pressure and the resultant hygroscopic white solid dissolved in methanol (25 cm^3) and *N*-methylmorpholine (1.12 cm^3, 10 mmol) added. The solvent was then removed under reduced pressure and the remaining residue dissolved in water and extracted with ethyl acetate (3 x 30 cm^3). The combined organic fractions were then dried ($MgSO_4$) and the solvent removed under reduced pressure to yield an off-white solid which was purified by silica chromatography using ethyl acetate/ methanol (95:5) as the eluent to give the product as a white solid (0.61 g, 62%), m.p. 108-10 °C; *m/z* (Found: *M*$^+$, 197.1159. Calc. for $C_9H_{15}N_3O_2$: 197.1163); $[\alpha]_D^{22}$ -82.6 (*c* 0.5 in MeOH); υ_{max} (Nujol)/cm^{-1} 3336 & 3215 (NH), 1655 (tertiary amide CO) and 1641 (secondary amide CO); δ_H (200 MHz; C^2HCl_3) 1.66 (3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.83-2.36 (4H, m, γCH_2 and βCH_2), 3.18 (3H, s, NCH_3), 3.76 (2H, m, δCH_2), 4.54 (1H, q, *J* 6.6, $C\underline{H}CH_3$) and 5.48 (1H, dd, J_1 4.6, J_2 8.0, αCH); δ_C (50.31 MHz; C^2HCl_3) 21.14 ($CH\underline{C}H_3$), 25.28 (γCH_2), 29.69 (βCH_2), 39.25 (NCH_3), 47.89 (δCH_2), 51.97 ($\underline{C}HCH_3$), 57.54 (αCH), 167.80 ($\underline{C}ONCH_3$) and 174.08 ($\underline{C}OCHCH_3$); *m/z* (EI) 197 (*M*$^+$, 15%), 182 (9, [*M* - CH_3]$^+$), 153 (17, [*M* - CH_3 - NCH_3]$^+$), 139 (5, [*M* - $NHNCH_3$ - CH_3 + H]$^+$), 127 (15, [*M* - $CHCH_3NHNCH_3$ + 2H]$^+$) and 70 (100, [C_4H_8N]$^+$).

(4*S*, 9a*S*)-2,4-Dimethyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (103)

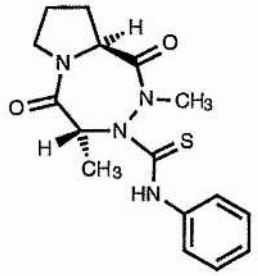
This compound was prepared in a manner identical with that for the (4*R*, 9a*S*) diastereomer (**102**), starting from *N-tert*-butoxycarbonyl-*N'*-(*N*-(2*R*)-chloropropanoyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**99**) (1.35 g, 5 mmol) instead of *N-tert*-butoxycarbonyl-*N'*-(*N*-(2*S*)-chloropropanoyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**98**) to give the product as a white solid (0.63 g, 64%), m.p. 112-15 °C; *m/z* (Found: M^+, 197.1159. Calc. for $C_9H_{15}N_3O_2$: 197.1163); $[\alpha]_D^{22}$ -119.4 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3299 & 3210 (NH), 1652 (tertiary amide CO) and 1634 (secondary amide CO); δ_H (200 MHz; C^2HCl_3) 1.57 (*B*, 3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.62 (*A*, 3H, d, *J* 6.8, $CHC\underline{H}_3$), 1.78-2.34 (*A* & *B*, 4H, m, γCH_2 and βCH_2), 3.14 (*A*, 3H, s, NCH_3), 3.18 (*B*, 3H, s, NCH_3), 3.59 (*A* & *B*, 1H, m, $\frac{1}{2}\delta CH_2$), 3.92 (*A* & *B*, 1H, m, $\frac{1}{2}\delta CH_2$), 4.48 (*A* & *B*, 1H, *J* 6.7, $C\underline{H}CH_3$), 5.40 (*A*, 1H, dd, J_1 4.4, J_2 7.8, αCH) and 5.42 (*B*, 1H, dd, J_1 4.4, J_2 8.6, αCH); δ_C (50.31 MHz; C^2HCl_3) 21.03 (*A*, $CH\underline{C}H_3$), 21.14 (*B*, $CH\underline{C}H_3$), 22.69 (*B*, γCH_2), 25.03 (*A*, γCH_2), 29.42 (*A*, βCH_2), 32.09 (*B*, βCH_2), 38.69 (*A*, NCH_3), 39.21 (*B*, NCH_3), 47.84 (*A* &*B*, δCH_2), 51.62 (*A*, $\underline{C}HCH_3$), 51.71 (*B*, $\underline{C}HCH_3$), 57.40 (*A*, αCH), 58.40 (*B*, αCH), 167.49 (*A*, $\underline{C}OCHCH_3$), 168.33 (*B*, $\underline{C}OCHCH_3$), 173.45 (*A*, $\underline{C}ONCH_3$) and 173.92 (*B*, $\underline{C}ONCH_3$); *m/z* (EI) 197 (M^+, 5%), 182 (19, $[M - CH_3]^+$), 153 (7, $[M - CH_3 - NCH_3]^+$), 139 (6, $[M - NHNCH_3 - CH_3 + H]^+$), 127 (23, $[M - CHCH_3NHNCH_3 + 2H]^+$), 97 (37, $[C_4H_8NCO]^+$) and 70 (100, $[C_4H_8N]^+$).

***N*-Propionyl-(2*S*)-proline methylamide (104)**

This compound was prepared in a manner identical with that for the *N*-benzyloxycarbonyl-(2*S*)-alanyl-(2*S*)-proline methylamide (**75**), starting from the *N*-propionyl-(2*S*)-proline methyl ester (0.93 g, 5 mmol) instead of *N*-benzyloxy-carbonyl-(2*S*)-alanyl-(2*S*)-proline methyl ester (**58**) to give the product as a waxy white solid (0.90 g, 98%), m.p. 49-52 °C; (Found: C, 58.65; H, 8.7; N, 15.35. Calc. for $C_9H_{16}N_2O_2$: C, 58.65; H, 8.75; N, 15.2%); $[\alpha]_D^{22}$ -83.1 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 3311 (NH), 1662 (secondary amide CO) and 1637 (tertiary amide CO); δ_H (200 MHz; C^2HCl_3) 1.08 (*c*, 3H, t, *J* 8.0, $CH_2C\underline{H}_3$), 1.11 (*t*, 3H, t, *J* 7.8, $CH_2C\underline{H}_3$), 1.66-2.48 (*t* & *c*, 4H, m, γCH_2 and βCH_2), 2.72 (*t*, 3H, d, *J* 4.8, NCH_3), 2.80 (*c*, 3H, d, *J* 4.8, NCH_3), 3.38 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 3.50 (*t* & *c*, 1H, m, $\frac{1}{2}\delta CH_2$), 4.29 (*c*, 1H, dd, J_1 5.2, J_2 5.8, αCH), 4.52 (*t*, 1H, dd, J_1 3.4, J_2 6.6, αCH), 6.57 (*c*, 1H, br, NH) and 7.18 (*t*, 1H, br, NH); δ_C (50.31 MHz; C^2HCl_3) 21.14 (*t*,

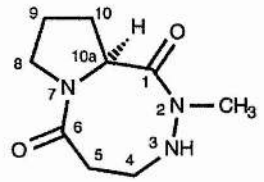
$CH_2\underline{C}H_3$), 8.77 (*c*, $CH_2\underline{C}H_3$), 22.52 (*c*, γCH_2), 24.56 (*t*, γCH_2), 25.82 (*t*, $\underline{C}H_2CH_3$), 26.02 (*c*, $\underline{C}H_2CH_3$), 27.14 (*c*, NCH_3), 27.43 (*t*, NCH_3), 28.61 (*t*, βCH_2), 31.98 (*c*, βCH_2), 46.55 (*c*, δCH_2), 47.12 (*t*, δCH_2), 59.71 (*t*, αCH), 60.61 (*c*, αCH), 172.49 (*t*, $\underline{C}ONCH_3$), 172.79 (*c*, $\underline{C}ONCH_3$), 173.03 (*c*, $\underline{C}OCH_2$) and 173.21 (*t*, $\underline{C}OCH_2$); *m/z* (EI) 184 (M^+, 10%), 171 (27, $[M - CH_3 + 2H]^+$), 154 (63, $[M - NHCH_3]^+$), 140 (55, $[M - NHCH_3 - CH_3 + H]^+$), 126 (82, $[M - CH_2CH_3 - NHCH_3 + H]^+$) and 70 (100, $[C_4H_8N]^+$).

(4*R*, 9a*S*)-2,4-Dimethyl-3-phenylthiocarbamyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (108)



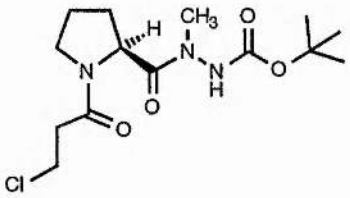
To a solution of (4*R*, 9a*S*)-2,4-dimethyl-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (**102**) (0.39 g, 2 mmol) in methanol (10 cm^3) was added phenyl *iso*-thiocyanate (1.35 g, 10 mmol). The resultant solution was stirred at 40 °C for 3 h and the solvent then removed under reduced pressure to give a yellow oil. The crude material was purified by silica column chromatography using ethyl acetate as the eluent to give the product as a white solid (0.51 g, 76%), m.p. 100-3 °C; *m/z* (Found: $[M + H]^+$, 333.1393. Calc. for $C_{16}H_{21}N_4O_2S$: 333.1385); $[\alpha]_D^{22}$ +29.2 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3210 (NH), 1675 (methylamide CO), 1638 (tertiary amide CO) and 1597 (thiourea CS); δ_H (200 MHz; C^2HCl_3) 1.66 (3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.83-2.36 (4H, m, γCH_2 and βCH_2), 3.18 (3H, s, NCH_3), 3.35 (1H, m, $\frac{1}{2}\delta CH_2$), 3.52 (1H, m, $\frac{1}{2}\delta CH_2$), 4.18 (1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.57 (1H, dd, $J_1 = J_2$ 7.0, αCH), 6.92 (3H, m, aromatic), 7.42 (2H, d, *J* 10.5, *ortho* aromatic) and 8.28 (1H, s, NH); δ_C (50.31 MHz; C^2HCl_3) 20.80 ($CH\underline{C}H_3$), 25.85 (γCH_2), 29.33 (βCH_2), 35.58 (NCH_3), 48.02 (δCH_2), 51.51 ($\underline{C}HCH_3$), 57.09 (αCH), 124.82 (*ortho* aromatic), 126.56 (*para* aromatic), 129.07 (*meta* aromatic), 138.88 (quat. aromatic), 168.43 ($\underline{C}ONCH_3$), 174.45 ($\underline{C}OCHCH_3$) and 182.07 (CS); *m/z* (CI) 333 ($[M + H]^+$, 15%), 245 (10, $[M - NHCS - CH_3 - CH_3 + 2H]^+$), 154 (6, $[M - PhNHCSN_2CH_3 + H]^+$), 137 (9, $[PhNHCHS]^+$), 127 (7, $[M - CHCH_3NHNCH_3 + 2H]^+$) and 94 (100, $[PhNH_3]^+$).

(10a*S*)-2-Methyl-2,3,4,5,8,9,10,10a-octahydropyrrolo-[2,1-d]-1,2,5-triazocine-1,6-dione (109)



This compound was prepared in a manner identical with that for the hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (**91**), using *N-tert*-butoxycarbonyl-*N'*-(*N*-3-chloropropanoyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**111**) (1.67 g, 5 mmol) instead of *N-tert*-butoxycarbonyl-*N'*-(*N*-chloroacetyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**89**) to give the product as a white solid (0.58 g, 60%), m.p. 129-32 °C; *m/z* (Found: $[M + H]^+$, 198.1241. Calc. for $C_9H_{16}N_3O_2$: 198.1243); $[\alpha]_D^{22}$ -93.2 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3380 (NH), 1655 (methylamide CO) and 1640 (tertiary amide CO); δ_H (200 MHz; C^2HCl_3) 1.73-2.30 (4H, m, γCH_2 and βCH_2), 2.79 (2H, t, *J* 7.0, $C\underline{H}_2NH$), 3.14 (3H, s, NCH_3), 3.49-3.98 (4H, m, δCH_2 and $COCH_2$), 3.94 (1H, br, NH) and 5.42 (1H, dd, J_1 4.0, J_2 7.8, αCH); δ_C (50.31 MHz; C^2HCl_3) 24.92 (γCH_2), 25.35 (βCH_2), 38.04 (CH_2NH), 39.61 (NCH_3), 39.97 ($CO\underline{C}H_2$), 48.07 (δCH_2), 57.13 (αCH), 168 82 ($\underline{C}ONCH_3$) and 174.22 ($\underline{C}OCH_2$); *m/z* (FAB) 198 ($[M + H]^+$, 53%), 170 (38, $[M - NCH_3 + 2H]^+$), 154 (37, $[M - NHNCH_3 + H]^+$), 142 (73, $[M - NCH_3NHCH_2 + 3H]^+$), 126 (54, $[M - NCH_3NHCH_2CH_2 + H]^+$) and 70 (100, $[C_4H_8N]^+$).

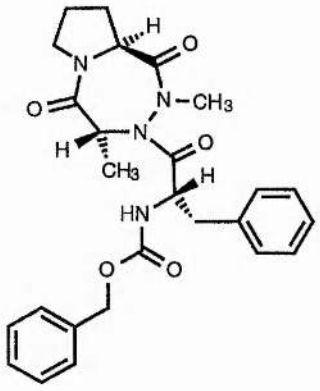
***N-tert*-Butoxycarbonyl-*N'*-(*N*-3-chloropropanoyl-(2*S*)-prolyl)-*N'*-methylhydrazine (111)**



This compound was prepared in a manner identical with that for *N-tert*-butoxycarbonyl-*N'*-(*N*-(2*S*)-chloropropanoyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**98**), using *N*-3-chloropropanoyl-(2*S*)-proline (**110**) (2.06 g, 10 mmol) to give the product as a white solid (2.93g, 88%), m.p. 108-11 °C; *m/z* (Found: $[M + H]^+$, 334.1537. Calc. for $C_{14}H_{25}{}^{35}ClN_3O_4$: 334.1536); $[\alpha]_D^{22}$ -21.7 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3221

(NH), 1733 (carbamate CO), 1684 (methylamide CO), 1632 (tertiary amide CO) and 665 (C-Cl); δ_H (200 MHz; C^2HCl_3) 1.44 (9H, s, $C(CH_3)_3$), 1.72-2.31 (4H, m, γCH_2 and βCH_2), 2.78 (2H, t, *J* 7.5, CH_2Cl), 3.05 (3H, s, NCH_3), 3.49 (1H, m, $\frac{1}{2}\delta CH_2$), 3.62 (1H, m, $\frac{1}{2}\delta CH_2$), 3.76 (2H, m, $COCH_2$), 4.97 (1H, dd, J_1 5.2, J_2 6.4, αCH) and 8.43 (1H, s, NH); δ_C (50.31 MHz; C^2HCl_3) 24.37 (γCH_2), 28.14 (βCH_2), 28.38 ($C(\underline{C}H_3)_3$), 34.87 (NCH_3), 47.59 (δCH_2), 56.98 (αCH), 81.48 ($\underline{C}(CH_3)_3$), 154.92 ($CO_2{}^tBu$), 168.58 ($\underline{C}ONCH_3$) and 173.71 ($\underline{C}OCH_2$); *m/z* (FAB) 334 ([*M* + H]$^+$, 36%), 298 (21, [*M* - Cl]$^+$), 288 (67, [*M* - CH_2Cl + 2H]$^+$), 278 (35, [*M* - $C(CH_3)_3$ + 2H]$^+$), 188 (68, [*M* - $NCH_3NHCO_2C(CH_3)_3$]$^+$), 70 (98, $[C_4H_8N]^+$) and 57 (100, $[C(CH_3)_3]^+$).

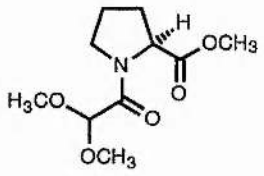
(4*R*, 9a*S*)-2,4-Dimethyl-3-(*N*-benzyloxycarbonyl-(2*S*)-phenylalanyl)-2,3,7,8,9,9a-hexahydropyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (114)



This compound was prepared in a manner identical with that for *N-tert*-butoxycarbonyl-*N'*-(*N*-chloroacetyl-(2*S*)-prolyl)-*N'*-methylhydrazine (**89**), using *N*-benzyloxycarbonyl-(2*S*)-phenylalanine (**113**) (1.50 g, 5 mmol) instead of *N*-chloroacetyl-(2*S*)-proline (**88**) and (4*R*, 9a*S*)-2,4-dimethyl-2,3,7,8,9,9a-hexahydro-pyrrolo-[2,1-d]-1,2,5-[6H]-triazepine-1,5-dione (**102**) (1.99 g, 5 mmol) instead of *N-tert*-butoxycarbonyl-*N'*-methylhydrazine (**87**). The crude material was purified by silica column chromatography using ethyl acetate/ light petroleum (7:3) as the eluent to give the product as a white solid (1.96 g, 82%), m.p. 82-5 °C; *m/z* (Found: [*M* + H]$^+$, 479.2282. Calc. for $C_{26}H_{31}N_4O_5$: 479.2294); $[\alpha]_D^{22}$ +0.8 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 3270 (NH), 1717 (carbamate CO), 1695 (methylamide CO), 1659 (secondary amide CO) and 1645 (tertiary amide CO); δ_H (200 MHz; C^2HCl_3) 1.59 (3H, d, *J* 6.6, $CHC\underline{H}_3$), 1.68-2.35 (4H, m, γCH_2 and βCH_2), 2.89 (3H, s, NCH_3), 3.09 (2H, d, *J* 7.6, $\frac{1}{2}CHC\underline{H}_2Ph$), 3.14 (2H, d, *J* 7.0, $\frac{1}{2}CHC\underline{H}_2Ph$), 3.56 (2H, m, δCH_2), 4.44 (1H, dd, J_1 7.4, J_2 8.8, αCH), 4.47 (1H, q, *J* 6.6, $C\underline{H}CH_3$), 4.70 (1H, m, $C\underline{H}NH$), 5.08

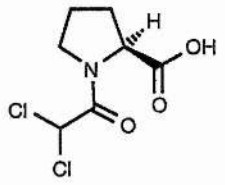
(2H, d, *J* 2.2, OCH_2Ph), 5.51 (1H, d, *J* 7.2, NH), 7.22 (5H, m, aromatic, $CHCH_2Ph$) and 7.32 (5H, s, aromatic, OCH_2Ph); δ_C (50.31 MHz; C^2HCl_3) 21.15 ($CH\underline{C}H_3$), 25.37 (γCH_2), 29.21 (βCH_2), 35.66 ($CH\underline{C}H_2Ph$), 38.66 (NCH_3), 47.90 (δCH_2), 51.99 ($\underline{C}HCH_3$), 55.46 (CHNH), 57.13 (αCH), 67.60 (OCH_2Ph), 127.75 (*para* aromatic), 128.51 (*ortho* aromatic) 128.72 (*para* aromatic), 129.07 (*ortho* aromatic), 129.30 (*meta* aromatic), 129.96 (*ortho* aromatic), 136.37 (quat. aromatic), 136.72 (quat. aromatic), 156.53 (CONH), 168.26 ($\underline{C}ONCH_3$), 170.90 ($\underline{C}OCHCH_3$) and 173.65 ($\underline{C}OCHCH_2Ph$); *m/z* (CI) 513 ([*M* + $2NH_3$ + H]$^+$, 100%), 479 (19, [*M* + H]$^+$), 403 (3, [*M* - Ph + H]$^+$), 345 (3, [*M* - CO_2CH_2Ph + 2H]$^+$), 295 (4, [*M* - OCH_2Ph - Ph + H]$^+$), 254 (5, [*M* - $NHCO_2CH_2Ph$ - Ph + 3H]$^+$) and 181 (32, [*M* - $COCHCH_2Ph$ - $NHCO_2CH_2Ph$ - CH_3]$^+$).

_N_-Dimethoxyacetyl-(2*S*)-proline methyl ester (118)



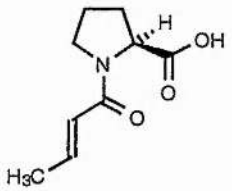
This compound was prepared in a manner identical with that for *N*-bromoacetyl-(2*S*)-proline methyl ester (**38**), using 2,2-dimethoxy-glyoxylic acid (**117**) (1.20 g, 10 mmol) instead of bromoacetic acid (**37**) to give a yellow oil which was purified by silica chromatography using ethyl acetate/ light petroleum (1:1) as the eluent to give the product as a clear oil (2.10 g, 91%), *m/z* (Found: [*M* + H]$^+$, 232.1181. Calc. for $C_{10}H_{18}NO_5$: 232.1184); $[\alpha]_D^{22}$ -73.7 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 1748 (ester CO), 1659 (amide CO) and 1438 (acetal C-O); δ_H (200 MHz; C^2HCl_3) 1.76-2.22 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.27 (*c*, 3H, s, $CHOC\underline{H}_3$), 3.32 (*c*, 3H, s, $CHOC\underline{H}_3$), 3.38 (*t*, 6H, s, $CH(OC\underline{H}_3)_2$), 3.47-3.76 (*t* & *c*, 2H, m, δCH_2), 3.66 (*t* & *c*, 3H, s, CO_2CH_3), 4.44 (*t*, 1H, dd, J_1 4.3, J_2 6.1, αCH), 4.62 (*c*, 1H, s, $C\underline{H}(OCH_3)_2$), 4.74 (*c*, 1H, dd, J_1 2.1, J_2 10.1, αCH) and 4.78 (*t*, 1H, s, $C\underline{H}(OCH_3)_2$); δ_C (50.31 MHz; C^2HCl_3) 21.49 (*c*, γCH_2), 25.18 (*t*, γCH_2), 28.73 (*t*, βCH_2), 31.33 (*c*, βCH_2), 46.62 (*t*, δCH_2), 46.84 (*c*, δCH_2), 52.12 (*t* & *c*, CO_2CH_3), 53.45 (*t*, $CHO\underline{C}H_3$), 54.07 (*c*, $CHO\underline{C}H_3$), 54.12 (*t*, $CHO\underline{C}H_3$), 54.53 (*c*, $CHO\underline{C}H_3$), 58.70 (*c*, αCH), 59.16 (*t*, αCH), 100.34 (*t*, $\underline{C}H(OCH_3)_2$), 102.28 (*c*, $\underline{C}H(OCH_3)_2$), 165.76 (*t*, $\underline{C}OCH(OCH_3)_2$), 165.87 (*c*, $\underline{C}OCH(OCH_3)_2$), 172.45 (*t*, $\underline{C}O_2CH_3$) and 172.96 (*c*, $\underline{C}O_2CH_3$); *m/z* (CI) 232 ([*M* + H]$^+$, 68%), 200 (100, [*M* - OCH_3]$^+$), 172 (49, [*M* - CO_2CH_3]$^+$), 128 (6, [*M* - $COCH(OCH_3)_2$]$^+$) and 75 (52, $[CH(OCH_3)_2]^+$).

_N_-Dichloroacetyl-(2_S_)-proline (122)



To a stirred suspension of (2*S*)-proline (**46**) (1.15 g, 10 mmol) in dry diethyl ether (100 cm^3) containing solid sodium hydrogen carbonate (5.0 g) was added dropwise dichloroacetyl chloride (1.48 g, 10 mmol) at 0 °C. The reaction mixture was stirred for 30 min at room temperature, filtered and the solvent removed under reduced pressure to give a white solid which was recrystallised from diethyl ether/ light petroleum to give the product as colourless crystals (1.43 g, 63%), m.p. 133-5 °C; (Found: C, 37.3; H, 3.9; N, 6.15. Calc. for $C_7H_9Cl_2NO_3$: C, 37.3 H, 4.05; N, 6.2%); *m/z* (Found: $[M + H]^+$, 226.0043. Calc. for $C_7H_{10}{}^{35}Cl_2NO_3$: 226.0038); $[\alpha]_D^{22}$ -84.3 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 1718 (acid CO), 1624 (amide CO) and 677 (C-Cl); δ_H (200 MHz; d_6-DMSO) 1.68-2.36 (*t* & *c*, 4H, m, βCH_2 and γCH_2), 3.31-3.80 (*t* & *c*, 2H, m, δCH_2), 4.28 (*t*, 1H, dd, J_1 3.8, J_2 8.8, αCH), 4.67 (*c*, 1H, dd, J_1 2.9, J_2 6.1, αCH), 6.72 (*c*, 1H, s, $CHCl_2$) and 6.91 (*t*, 1H, s, $CHCl_2$); δ_C (50.31 MHz; d_6-DMSO) 21.84 (*c*, γCH_2), 24.72 (*t*, γCH_2), 28.85 (*t*, βCH_2), 30.63 (*c*, βCH_2), 47.11 (*t* & *c*, δCH_2), 59.41 (*c*, αCH), 59.80 (*t*, αCH), 66.18 (*c*, $CHCl_2$), 66.43 (*t*, $CHCl_2$), 161.65 (*t*, $\underline{C}OCHCl_2$), 162.47 (*c*, $\underline{C}OCHCl_2$) and 172.65 (*t* & *c*, CO_2H); *m/z* (CI) 226 ($[M + H]^+$, 100%), 208 (7, $[M - OH]^+$), 190 (13, $[M - Cl]^+$), 180 (14, $[M - CO_2H]^+$), 156 (7, $[M - 2Cl + H]^+$), 112 (16, $[COCHCl_2]^+$ and 70 (43, $[C_4H_8N]^+$).

_N_-Crotonyl-(2S)-proline (125)



To a vigorously stirred solution of (2*S*)-proline (**46**) (1.27 g, 11 mmol) in sodium hydroxide solution (80 cm^3, 5% w/v) was added dropwise *trans*-crotonyl chloride (**124**) (1.05 g, 10 mmol) at 0 °C. The reaction mixture was allowed to warm to room temperature and stirred for 1 h, before being acidified to pH 1 with concentrated HCl.

The resulted solution was extracted with ethyl acetate (3 x 30 cm^3), dried ($MgSO_4$) and the solvent removed under reduced pressureto give a white solid. Recrystallisation from ethyl acetate gave the product as colourless crystals (1.17 g, 64%), m.p. 161-3 °C; (Found: C, 59.2; H, 7.2; N, 7.55; M^+, 183.0893. Calc. for $C_9H_{13}NO_3$: C, 59.0; H, 7.15; N, 7.65%; M^+, 183.0895); $[\alpha]_D^{22}$ +99.3 (*c* 1.0 in MeOH); υ_{max} (Nujol)/cm^{-1} 1713 (acid CO), 1659 (amide CO) and 1583 (C=C); δ_H (200 MHz; C^2HCl_3) 1.93 (3H, dd, J_1 1.6, J_2 7.1, CHC<u>H</u>$_3$), 1.95-2.12 (3H, m, γCH_2 and $\frac{1}{2}\beta CH_2$), 2.10-2.45 (1H, m, $\frac{1}{2}\beta CH_2$), 3.45-3.85 (2H, m, δCH_2), 4.47 (1H, dd, J_1 3.3, J_2 7.9, αCH), 6.34 (1H, dd, J_1 1.6, J_2 15.0, C<u>H</u>CHCH$_3$), 6.87 (1H, dd, J_1 7.1, J_2 15.0, C<u>H</u>CH$_3$) and 10.4 (1H, s, CO_2H); δ_C (50.31 MHz; C^2HCl_3) 18.60 (CH<u>C</u>H$_3$), 26.02 (γCH_2), 30.61 (βCH_2), 48.51 (δCH_2), 60.67 (αCH), 124.03 (<u>C</u>HCHCH$_3$), 143.89 (<u>C</u>HCH$_3$), 167.38 (N<u>C</u>O) and 175.91 (CO_2H); *m/z* (EI) 183 (M^+, 3%), 138 (40, [M - CO_2H]$^+$), 124 (39, [M - CH_3CH=CH - OH + H]$^+$), 96 (19, [C_4H_7NCO]$^+$) and 70 (100, [C_4H_8N]$^+$).

N-Crotonyl-(2*S*)-proline methyl ester (126)

A boron trifloride methanol (BF_3.MeOH) complex (22 cm^3, 100 mmol) was added to *N*-crotonyl-(2*S*)-proline (**125**) (3.66 g, 20 mmol) and the resultant solution refluxed for 2 h, and cooled before the addition of a saturated solution of $NaHCO_3$ (82 cm^3). The reaction mixture was then extracted with diethyl ether (3 x 50 cm^3), dried ($MgSO_4$) and the solvent removed under reduced pressure to give the product as a clear oil (2.87 g, 73%), *m/z* (Found: [M + H]$^+$, 198.1127. Calc. for $C_{10}H_{16}NO_3$: 198.1130); $[\alpha]_D^{22}$ -113.7 (*c* 1.0 in MeOH); υ_{max} (thin film)/cm^{-1} 1746 (ester CO), 1666 (amide CO) and 1620 (C=C); δ_H (200 MHz; C^2HCl_3) 1.82 (3H, dd, J_1 1.6, J_2 7.0, CHC<u>H</u>$_3$), 1.85-2.22 (4H, m, γCH_2 and βCH_2), 3.48-3.61 (2H, m, δCH_2), 3.67 (3H, s, OCH_3), 4.48 (1H, dd, J_1 3.5, J_2 7.6, αCH), 6.09 (1H, dd, J_1 1.6, J_2 15.0, C<u>H</u>CHCH$_3$) and 6.90 (1H, dq, J_1 7.0, J_2 15.0, C<u>H</u>CH$_3$); δ_C (50.3 MHz; C^2HCl_3) 18.45 (CH<u>C</u>H$_3$), 25.17 (γCH_2), 29.48 (βCH_2), 47.19 (δCH_2), 52.47 (OCH_3), 59.13 (αCH), 122.77 (<u>C</u>HCHCH$_3$), 142.42 (<u>C</u>HCH$_3$), 164.16 (NCO) and 173.22 (<u>C</u>O$_2$CH$_3$); *m/z* (CI) 198 ([M + H]$^+$, 100%), 184 (13, [M - CH_3 + 2H]$^+$), 166 (18, [M - OCH_3]$^+$), 138 (27, [M - CO_2CH_3]$^+$) and 70 (7, [C_4H_8N]$^+$).

Appendix 1

Solid-phase peptide synthesis

The peptides were prepared using a Rainin PS3 automated solid phase peptide synthesizer. The resin (Wang-type) and all amino acids (Fmoc protected) were purchased from Novabiochem Ltd. PyBOP was used as the coupling reagent and was also purchased from Novabiochem. Piperazine in DMF (20%, v/v) was used as the deprotection solution and NMM in DMF (0.4 mol dm^{-3}) was used as the activation solution. The peptides were prepared on a 0.1 mmol scale with coupling times of 60 min used throughout the syntheses. Double couplings were carried out for difficult amino acids (Pro, Val, Asn).

The peptides were cleaved from the resin using a mixture of TFA, water and anisole (95: 2.5: 2.5 respectively) and left for 6 h at 25 °C. To ensure complete removal of non-polar contaminants, the peptides were dissolved in water and acetic acid (20 cm^3, 9:1) and washed with ethyl acetate (2 x 10 cm^3), CH_2Cl_2 (10 cm^3) and diethyl ether (10 cm^3). The aqueous phase was then lyophilized under high vacuum and the dry peptides stored at -20 °C.

The purity of the peptides was checked by HPLC. The dry peptide (1 mg) was dissolved in water (1 cm^3), filtered and injected (10 µl) onto a C_{18} column equilibrated with water containing TFA (0.1%). The peptides were eluted with 0-60% acetonitrile over 30 min and monitored at 214 nm.

Appendix 2

NMR and CD analysis of oligopeptides

NMR spectra were recorded on a Varian VXR6005 600 MHz spectrometer at the Edinburgh University Ultra-high field NMR centre. All experiments were performed on solutions of 10 mg of peptide in 1 cm^3 of solvent.

The CD experiments were carried out by Dr. Nick Price and Ms. Sharon Kelly at the BBSRC service centre in the department of biomolecular sciences, Stirling University. The machine used was a JASCO J-600 spectropolarimeter and the cell path length was 0.02 cm. All experiments were performed at 37 °C on solutions of 0.5 mg of peptide in 1 cm^3 of solvent.

Appendix 3

Crystallography data

Table 3.1: *Bond lengths (Å) and angles (°) for (48)*

Br-C(2)	1.963(6)	N-C(4)	1.458(8)
O(1)-C(3)	1.211(7)	C(1)-C(2)	1.485(10)
O(2)-C(8)	1.175(8)	C(2)-C(3)	1.529(8)
O(3)-C(8)	1.334(7)	C(4)-C(5)	1.500(10)
O(3)-C(9)	1.428(8)	C(5)-C(6)	1.443(12)
N-C(3)	1.346(7)	C(6)-C(7)	1.518(9)
N-C(7)	1.453(7)	C(7)-C(8)	1.521(9)
C(8)-O(3)-C(9)	116.1(5)	N-C(4)-C(5)	102.0(5)
C(3)-N-C(7)	117.8(5)	C(6)-C(5)-C(4)	108.4(6)
C(3)-N-C(4)	128.3(5)	C(5)-C(6)-C(7)	105.9(6)
C(7)-N-C(4)	113.1(5)	N-C(7)-C(6)	103.7(5)
C(1)-C(2)-C(3)	113.9(6)	N-C(7)-C(8)	111.7(5)
C(1)-C(2)-Br	110.0(5)	C(6)-C(7)-C(8)	112.4(6)
C(3)-C(2)-Br	107.1(4)	O(2)-C(8)-O(3)	124.3(6)
O(1)-C(3)-N	121.5(5)	O(2)-C(8)-C(7)	126.0(6)
O(1)-C(3)-C(2)	120.2(5)	O(3)-C(8)-C(7)	109.6(5)
N-C(3)-C(2)	118.3(5)		

Table 3.2: *Bond lengths (Å) and angles (°) for (81)*

O(1)-C(1)	1.318(2)	C(2)-C(3)	1.538(2)
O(2)-C(1)	1.210(2)	C(3)-C(4)	1.521(3)
O(3)-C(6)	1.251(2)	C(4)-C(5)	1.524(2)
N-C(6)	1.334(2)	C(6)-C(7)	1.478(2)
N-C(2)	1.462(2)	C(7)-C(8)	1.312(2)
N-C(5)	1.478(2)	C(8)-C(9)	1.492(2)
C(1)-C(2)	1.522(2)		
C(6)-N-C(2)	120.28(11)	C(4)-C(3)-C(2)	103.95(12)
C(6)-N-C(5)	127.63(11)	C(3)-C(4)-C(5)	103.58(13)
C(2)-N-C(5)	112.08(11)	N-C(5)-C(4)	102.99(12)
O(2)-C(1)-O(1)	124.52(13)	O(3)-C(6)-N	119.47(12)
O(2)-C(1)-C(2)	121.43(13)	O(3)-C(6)-C(7)	122.31(12)
O(1)-C(1)-C(2)	113.94(12)	N-C(6)-C(7)	118.22(12)
N-C(2)-C(1)	113.59(11)	C(8)-C(7)-C(6)	122.08(13)
N-C(2)-C(3)	103.80(11)	C(7)-C(8)-C(9)	125.9(2)
C(1)-C(2)-C(3)	110.88(12)		

Table 3.3: *Bond lengths (Å) and angles (°) for* ***(82)***

O(01)-C(01)	1.224(2)	C(03)-C(04B)	1.433(5)	N(2)-N(3)	1.421(2)
O(02)-C(06)	1.238(2)	C(04A)-C(05)	1.460(5)	N(2)-C(7)	1.455(3)
N(01)-C(02)	1.491(2)	C(04B)-C(05)	1.535(5)	N(2)-C(8)	1.456(3)
N(01)-C(05)	1.476(2)	C(06)-C(07)	1.510(3)	N(3)-C(1)	1.326(3)
N(01)-C(06)	1.330(2)	O(02)-H(3)	2.01(3)	C(1)-C(2)	1.520(3)
N(02)-N(03)	1.428(2)	N(03)-H(03)	0.87(2)	C(2)-C(3)	1.524(3)
N(02)-C(07)	1.449(2)	O(1)-C(1)	1.224(2)	C(3)-C(4)	1.517(3)
N(02)-C(08)	1.460(2)	O(2)-C(6)	1.243(2)	C(4)-C(5)	1.507(3)
N(03)-C(01)	1.336(2)	N(1)-C(2)	1.490(2)	C(6)-C(7)	1.508(3)
C(01)-C(02)	1.522(3)	N(1)-C(5)	1.471(3)	O(2)-H(03)	2.02(2)
C(02)-C(03)	1.517(3)	N(1)-C(6)	1.329(2)	N(3)-H(3)	0.89(3)
C(03)-C(04A)	1.529(5)				

C(06)-O(02)-H(3)	128.8(7)	C(6)-O(2)-H(03)	130.3(6)
C(06)-N(01)-C(05)	120.1(2)	C(6)-N(1)-C(5)	120.3(2)
C(06)-N(01)-C(02)	128.8(2)	C(6)-N(1)-C(2)	129.2(2)
C(05)-N(01)-C(02)	110.73(15)	C(5)-N(1)-C(2)	110.6(2)
N(03)-N(02)-C(07)	110.23(14)	N(3)-N(2)-C(7)	109.74(14)
N(03)-N(02)-C(08)	109.1(2)	N(3)-N(2)-C(8)	109.0(2)
C(07)-N(02)-C(08)	112.85(15)	C(7)-N(2)-C(8)	113.6(2)
H(03)-N(03)-C(01)	121.6(14)	H(3)-N(3)-C(1)	120.6(16)
H(03)-N(03)-N(02)	113.7(14)	H(3)-N(3)-N(2)	119.9(16)
C(01)-N(03)-N(02)	118.2(2)	C(1)-N(3)-N(2)	119.2(2)
O(01)-C(01)-N(03)	122.1(2)	O(1)-C(1)-N(3)	122.7(2)
O(01)-C(01)-C(02)	123.5(2)	O(1)-C(1)-C(2)	122.7(2)
N(03)-C(01)-C(02)	114.4(2)	N(3)-C(1)-C(2)	114.5(2)
N(01)-C(02)-C(03)	103.7(2)	N(1)-C(2)-C(1)	111.33(14)
N(01)-C(02)-C(01)	108.84(14)	N(1)-C(2)-C(3)	103.49(15)
C(03)-C(02)-C(01)	113.8(2)	C(1)-C(2)-C(3)	113.5(2)
C(02)-C(03)-C(04A)	105.9(2)	C(4)-C(3)-C(2)	103.9(2)
C(02)-C(03)-C(04B)	106.0(2)	C(5)-C(4)-C(3)	103.7(2)
C(03)-C(04A)-C(05)	105.2(3)	N(1)-C(5)-C(4)	104.4(2)
C(03)-C(04B)-C(05)	106.2(3)	O(2)-C(6)-N(1)	120.1(2)
N(01)-C(05)-C(04A)	105.5(2)	O(2)-C(6)-C(7)	117.8(2)
N(01)-C(05)-C(04B)	100.9(2)	N(1)-C(6)-C(7)	122.1(2)
O(02)-C(06)-N(01)	120.8(2)	N(2)-C(7)-C(6)	114.7(2)
O(02)-C(06)-C(07)	117.3(2)		
N(01)-C(06)-C(07)	121.9(2)		
N(02)-C(07)-C(06)	115.0(2)		

Table 3.4: *Bond lengths (Å) and angles (°) for* ***(117)***

O(1)-C(1)	1.216(3)	N(3)-C(1)	1.338(2)
O(2)-C(6)	1.235(2)	N(3)-H(3)	0.89(2)
N(1)-C(6)	1.337(2)	C(1)-C(2)	1.529(3)
N(1)-C(5)	1.473(2)	C(2)-C(3)	1.529(3)
N(1)-C(2)	1.484(2)	C(3)-C(4)	1.474(4)
N(2)-N(3)	1.422(2)	C(4)-C(5)	1.492(4)
N(2)-C(7)	1.461(2)	C(6)-C(7)	1.526(2)
N(2)-C(9)	1.466(2)	C(7)-C(8)	1.527(3)

C(6)-N(1)-C(5)	120.49(15)	N(1)-C(2)-C(1)	107.42(14)
C(6)-N(1)-C(2)	128.22(14)	N(1)-C(2)-C(3)	103.0(2)
C(5)-N(1)-C(2)	110.90(14)	C(1)-C(2)-C(3)	113.9(2)
N(3)-N(2)-C(7)	110.70(13)	C(4)-C(3)-C(2)	106.1(2)
N(3)-N(2)-C(9)	107.6(2)	C(3)-C(4)-C(5)	105.2(2)
C(7)-N(2)-C(9)	114.1(2)	N(1)-C(5)-C(4)	103.6(2)
C(1)-N(3)-N(2)	119.1(2)	O(2)-C(6)-N(1)	120.8(2)
C(1)-N(3)-H(3)	123.4(15)	O(2)-C(6)-C(7)	117.2(2)
N(2)-N(3)-H(3)	117.5(15)	N(1)-C(6)-C(7)	121.91(15)
O(1)-C(1)-N(3)	122.6(2)	N(2)-C(7)-C(6)	113.99(13)
O(1)-C(1)-C(2)	124.4(2)	N(2)-C(7)-C(8)	109.74(15)
N(3)-C(1)-C(2)	112.9(2)	C(6)-C(7)-C(8)	105.9(2)

REFERENCES

References

1. C. U. T. Hellen and E. Wimmer, *Curr. Opin. Biotech.*, 1992, **3**, 643-649.

2. W. G. Dougherty and B. L. Semler, *Microbiol. Rev.*, 1993, **57**, 781-822.

3. G. Stanway, *J. Gen. Virol.*, 1990, **71**, 2483-2501.

4. C. U. T. Hellen and E. Wimmer, *Experientia*, 1992, **48**, 201-215.

5. M. G. Rossmann, E. Arnold, J. W. Erickson, G. Kramer, M. Luo, A. G. Mosser, R. R. Rueckert, B. Sherry and G. Vriend, *Nature*, 1985, **317**, 145-153.

6. R. Acharya, E. Fry, D. Stuart, G. Fox, D. Rowlands and F. Brown, *Nature*, 1989, **337**, 709-716.

7. A. Aguilera, *Interferon y Biotecnologia*, 1987, **4**, 284-285.

8. F. Brown, *Vaccine*, 1988, **6**, 180-182.

9. N. Parry, G. Fox, D. Rowlands, F. Brown, E. Fry, R. Acharya, D. Logan and D. Stuart, *Nature*, 1990, **347**, 569-572.

10. R. Dimarchi, G. Brooke, C. Gale, V. Cracknell, T. Doel and N. Mowat, *Science*, 1986, **232**, 639-641.

11. A. A. M. Thomas, R. J. Woortmeijer, W. Puijk and S. J. Barteling, *J. Virol.*, 1988, **62**, 2782-2789.

12. Q. C. Xie, D. McCahon, J. R. Crowther, G. J. Belsham and K. C. McCullough, *J. Gen. Virol.*, 1987, **68**, 1637-1647.

13. D. Logan, R. Abu-Ghazaleh, W. Blakemore, S. Curry, T. Jackson, A. King, S. Lea, R. Lewis, J. Newman, N. Parry, D. Rowlands, D. Stuart and E. Fry, *Nature*, 1993, **362**, 566-568.

14. G. J. Belsham, *Prog. Biophys. Mol. Biol.*, 1993, **60**, 241-260.

15. E. Fry, D. Logan, R. Acharya, G. Fox, D. Rowlands, F. Brown and D. Stewart, *Sem. Virol.*, 1990, **1**, 439-451.

16. J. L. Bittle, R. A. Houghton, H. Alexander, T. M. Shinnick, J. G. Sutcliffe, R. A. Lerner, D. J. Rowlands and F. Brown, *Nature*, 1982, **298**, 30-33.

17. T. R. Doel, C. Gale, G. Brooke and R. DiMarchi, *J. Gen. Virol.*, 1988, **69**, 2403-2406.

18. M. J. Francis, G. Z. Hastings, A. D. Syred, B. McGinn, F. Brown and D. J. Rowlands, *Nature*, 1987, **300**, 168-170.

19. S. A. Lewis, D. O. Morgan and M. J. Grubman, *J. Virol.*, 1991, **65**, 6572-6580.

20. G. J. Belsham, J. K. Brangwyn, M. D. Ryan, C. C. Abrams and A. M. Q. King, *Virology*, 1990, **176**, 524-530.

21. L. G. Kleina and M. J. Grubman, *J. Virol.*, 1992, **66**, 7168-7175.

22. M. A. Devaney, V. N. Vakharia, R. E. Lloyd, E. Ehrenfeld and M. J. Grubman, *J. Virol.*, 1988, **62**, 4407-4409.

23. J. J. Harber, J. Bradley, C. W. Anderson and E. Wimmer, *J. Virol.*, 1991, **65**, 326-334.

24. E. Arnold, M. Luo, G. Vriend, M. G. Rossmann, A. C. Palmenberg, G. D. Parks, M. J. H. Nicklin and E. Wimmer, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 21-25.

25. D. Marc, M. Girard and S. Van der Werf, *J. Gen.Virol.*, 1991, **72**, 1151-1157.

26. C. U. T. Hellen and E. Wimmer, *Virology*, 1992, **187**, 391-397.

27. P. Talbot, D. J. Rowlands, J. N. Burroughs, D. V. Sangar and F. Brown, *J. Gen. Virol.*, 1973, **19**, 369-380.

28. B. E. Clarke and D. V. Sangar, *J. Gen. Virol.*, 1988, **69**, 2313-2325.

29. B. H. Robertson, M. J. Grubman, G. N. Weddell, D. M. Moore, J. D. Welsh, T. Fischer, D. J. Dowbenko, D. G. Yansura, B. Small and Dennis G. Kleid, *J. Virol.*, 1985, **54**, 651-660.

30. S. Forss, K. Strebel, E. Beck and H. Schaller, *Nucleic Acids Res.*, 1984, **12**, 6587-6601.

31. M. A. Lawson and B. L. Semler, *Curr. Topics Microbiol. Immunol.*, 1990, **161**, 155-188.

32. A. C. Palmenberg, *Ann. Rev. Microbiol.*, 1990, **44**, 603-623.

33. M. D. Ryan, G. J. Belsham and A. M. Q. King, *Virology*, 1989, **173**, 35-45.

34. V. N. Vakharia, M. A. Devaney, D. M. Moore, J. J. Dunn and M. J. Grubman, *J. Virol.*, 1987, **61**, 3199-3207.

35. B. E. Clarke, D. V. Sangar, J. N. Burroughs, S. E. Newton, A. R. Carroll and D. J. Rowlands, *J. Gen. Virol.*, 1985, **66**, 2615-2626.

36. M. D. Ryan, A. M. Q. King and G. P. Thomas, *J. Gen. Virol.*, 1991, **72**, 2727-2732.

37. A. C. Palmenberg and R. R. Rueckert, *J. Virol.*, 1982, **41**, 244-249.

38. K. Strebel and E. Beck, *J. Virol.*, 1986, **58**, 893-899.

39. C. Mirzayan, R. Ingraham and E. Wimmer, *J. Gen. Virol.*, 1991, **137**, 1159-1163.

40. W. S. Blair and B. L. Semler, *Curr. Opin. Cell Biol.*, 1991, **3**, 1039-1045.

41. S. F. Yu and R. E. Lloyd, *Virology*, 1991, **182**, 615-625.

42. S. F. Yu and R. E. Lloyd, *Virology*, 1992, **186**, 725-735.

43. A. C. Palmenberg, G. D. Parks, D. J. Hall, R. H. Ingraham, T. W. Seng and P. V. Pallai, *Virology*, 1992, **190**, 754-762.

44. M. D. Ryan and J. Drew, *EMBO J.*, 1994, **13**, 928-933.

45 N. Percy, W. S. Barclay, A. Garcia-Sastre and P. Palese, *J. Virol.*, 1994, **68**, 4486-4492.

46. F. D. Sönnichsen, J. E. Van Eyk, R. S. Hodges and B. D. Sykes, *Biochemistry*, 1992, **31**, 8790-8798.

47. A. Bierzynski, P. S. Kim and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 2470-2474.

48. G. Siligardi, A. F. Drake, P. Mascagni, P. Neri, L. Lozzi, N. Niccolai and W. A. Gibbons, *Biochem. Biophys. Res. Comm.*, 1987, **143**, 1005-1011.

49. M. D. Bruch and L. M. Gierasch, *J. Biol. Chem.*, 1990, **265**, 3851-3858.

50. D. Marion, M. Zasloff and A. Bax, *FEBS Lett.*, 1988, **227**, 21-26.

51. D. Greff, S. Fermandjian, P. Fromageot, M. C. Khosla, R. R. Smeby and F. M. Bumpus, *Eur. J. Biochem.*, 1976, **61**, 297-305.

52. J. W. Nelson and N. R. Kallenbach, *Biochemistry,* 1989, **28**, 5256-5261.

53. Y. Yamamoto, T. Ohkubo, A. Kohara, T. Tanaka, T. Tanaka and M. Kikuchi, *Biochemistry,* 1990, **29**, 8998-9006.

54. W. Braun, G. Wider, K. H. Lee and K. Wüthrich, *J. Mol. Biol.,* 1983, **169**, 921-948.

55. K. Wüthrich, M. Billeter and W. Braun, *J. Mol. Biol.,* 1984, **180**, 715-740.

56. M. P. Williamson, D. Marion and K. Wüthrich, *J. Mol. Biol.,* 1984, **173**, 341-359.

57. C. Burke, K. H. Mayo, A. P. N. Skubitz and L. T. Furcht, *J. Biol. Chem.,* 1991, **266**, 19407-19412.

58. E. R. P. Zuiderweg, R. Kaptein and K. Wüthrich, *Proc. Natl. Acad. Sci. USA,* 1983, **80**, 5837-5841.

59. S. W. Provencher and J. Glöckner, *Biochemistry,* 1981, **20**, 33-37.

60. D. Wemmer and N. R. Kallenbach, *Biochemistry,* 1983, **22**, 1901-1906.

61. A. L. Breeze, T. S. Harvey, R. Bazzo and I. D. Campbell, *Biochemistry,* 1991, **30**, 575-582.

62. F. Macquaire, F. Baleux, E. Giaccobi, T. Huynh-Dinh, J. -M. Neumann and A. Sanson, *Biochemistry,* 1992, **31**, 2576-2582.

63. M. J. Scanlon and R. S. Norton, *Protein Science,* 1994, **3**, 1121-1124.

64. M. P. Williamson, M. J. Hall and B. K. Handa, *Eur. J. Biochem.,* 1986, **158**, 527-536.

65. M. D. Ryan and J. Drew, *EMBO J.*, 1994, **13**, 928-933.

66. N. Percy, W. S. Barclay, A. Garcia-Sastre and P. Palese, *J. Virol.,* 1994, **68**, 4486-4492.

67. D. Gani, M. D. Ryan and J. Wilkie, unpublished results.

68. A. Yaron and F. Naider, *Crit. Rev. Biochem. Mol. Biol.*, 1993, **28**, 31-81.

69. C. M. Deber, M. Glibowicka and G. A. Woolley, *Biopolymers,* 1990, **29**, 149-157.

70. M. W. MacArthur and J. M. Thornton, *J. Mol. Biol.*, 1991, **218**, 397-412.

71. M. P. Williamson, *Biochem. J.,* 1994, **297**, 249-260.

72. J. B. Dame, J. L. Williams, T. F. M^{c}Cutchan, J. L. Weber, R. A. Wirtz, W. T. Hockmeyer, W. L. Maloy, J. D. Haynes, I. Schneider, D. Roberts, G. S. Sanders, E. P. Reddy, C. L. Diggs and L. H. Miller, *Science*, 1984, **225**, 593-599.

73. P. Lawton, C. Barrault, P. Grellier, J. Schrevel and H. -G. Heidrich, *Parasitol. Res.,* 1991, **77**, 720-722.

74. R. Baer, A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Séguin, P. S. Tufnell and B. G. Barrell, *Nature,* 1984, **310**, 207-211.

75. D. E. Stewart, A. Sarkar and J. E. Wampler, *J. Mol. Biol.,* 1990, **214**, 253-260.

76. P. N. Lewis, F. A. Momany and H. A. Scheraga, *Biochim. Biophys. Acta,* 1973, **303**, 211-229.

77. C. M. Wilmot and J. M. Thornton, *J. Mol. Biol.,* 1988, **203**, 221-232.

78. G. Müller, M. Gurrath, M. Kurz and H. Kessler, *Proteins: Structure, Function and Genetics,* 1993, **15**, 235-251.

79. H. A. Nagarajaram, P. K. C. Paul, K. Ramanarayanan, K. V. Soman and C. Ramakrishnan, *Int. J. Peptide Protein Res.*,1992, **40**, 383-394.

80. H. Kessler, H. Matter, G. Gemmecker, H. -J. Diehl, C. Isernia and S. Mronga, *Int. J. Peptide Protein Res.*, 1994, **43**, 47-61.

81. W. S. Blair and B. L. Semler, *Curr. Opin. Cell Biol.,* 1991, **3**, 1039-1045.

82. J. F. Brandts, H. R. Halvorson and M. Brennan, *Biochemistry*, 1975, **14**, 4953-4963.

83. K. Wüthrich and C. Grathwohl, *FEBS Lett.,* 1974, **43**, 337-340.

84. G. Kartha, T. Ashida and M. Kakudo, *Acta Cryst. B,* 1974, **30**, 1861-1866.

85. R. Freedman, *Nature*, **279**, 756-757.

86. M. A. Phillips and R. J. Fletterick, *Curr. Opin. Struct. Biol.,* 1992, **2**, 713-720.

87. H. Neurath, *Fed. Proc.,* 1985, **44**, 2907-2913.

88. N. Resnick and M. A. Zasloff, *Curr. Opin. Cell Biol.,* 1992, **4**, 1032-1036.

89. P. A. Evans, C. M. Dobson, R. A. Kautz, G. Hatfull and R. O. Fox, *Nature*, 1987, **329**, 266-268.

90. C. M. Dobson, P. A. Evans and K. L. Williamson, *FEBS Lett.,* 1984, **168**, 331-334.

91. W. J. Chazin, J. Kördel, T. Drakenberg, P. Brodin, T. Grundström and S. Forsén, *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 2195-2198.

92. F. J. Luque and M. Orozco, *J. Chem. Soc., Perkin Trans. 2,* 1993, 683-690.

93. F. J. Luque and M. Orozco, *J. Org. Chem.,* 1993, **58**, 6397-6405.

94. D. M. Schnur, Y. H. Yuh and D. R. Dalton, *J. Org. Chem.,* 1989, **54**, 3779-3785.

95. E. S. Eberhardt, S. N. Loh and R. T. Raines, *Tetrahedron Lett.*, 1993, **34**, 3055-3056.

96. H. N. Cheng and F. A. Bovey, *Biopolymers,* 1977, **16**, 1465-1472.

97. C. A. Evans and D. L. Rabenstein, *J. Am. Chem. Soc.*, 1974, **96**, 7312-7317.

98. E. S. Eberhardt, S. N. Loh, A. P. Hinck and R. T. Raines, *J. Am. Chem. Soc.*, 1992, **114**, 5437-5439.

99. K. Luthman and U. Hacksell, *Acta Chem. Scand.,* 1993, **47**, 461-468.

100. C. K. Larive and D. L. Rabenstein, *J. Am. Chem. Soc.*, 1993, **115**, 2833-2836.

101. R. E. London, N. A. Matwiyoff, J. M. Stewart and J. R. Cann, *Biochemistry,* 1978, **17**, 2277-2283.

102. R. K. Harrison and R. L. Stein, *J. Am. Chem. Soc.*, 1992, **114**, 3464-3471.

103. R. E. London, J. M. Stewart, R. Williams, J. R. Cann and N. A. Matwiyoff, *J. Am. Chem. Soc.*, 1979, **101**, 2455-2462.

104. M. J. O. Anteunis, F. A. M. Borremans, J. M. Stewart and R. E. London, *J. Am. Chem. Soc.*, 1981, **103**, 2187-2191.

105. R. E. London, J. M. Stewart and J. R. Cann, *Biochem. Pharm.*, 1990, **40**, 41-48.

106. C. Vicent, S. C. Hirst, F. Garcia-Tellado and A. D. Hamilton, *J. Am. Chem. Soc.*, 1991, **113**, 5466-5467.

107. D. E. Stewart, A. Sarkar and J. E. Wampler, *J. Mol. Biol.*, 1990, **214**, 253-260.

108. B. Vitoux, A. Aubry, M. T. Cung and M. Marraud, *Int. J. Peptide Protein Res.*, 1986, **27**, 617-632.

109. A. L. Love, T. D. Alger and R. K. Olsen, *J. Phys. Chem.*, 1972, **76**, 853-855.

110. D. F. Mierke, T. Yamazaki, O. E. Said-Nejad, E. R. Felder and M. Goodman, *J. Am. Chem. Soc.*, 1989, **111**, 6847-6849.

111. H. Kessler, U. Anders and M. Schudok, *J. Am. Chem. Soc.*, 1990, **112**, 5908-5916.

112. D. K. Sukumaran, M. Prorok and D. S. Lawrence, *J. Am. Chem. Soc.*, 1991, **113**, 706-707.

113. E. Bairaktari, D. F. Mierke, S. Mammi and E. Peggion, *J. Am. Chem. Soc.*, 1990, **112**, 5383.

114. F. X. Schmid, *Ann. Rev. Biophys. Biomol. Struct.*, 1993, **22**, 123-143.

115. G. Fischer, H. Bang and C. Mech, *Biomed. Biochim. Acta*, 1984, **43**, 1101-1111.

116. K. Lang and F. X. Schmid, *Nature*, 1988, **331**, 453-455.

117. H. P. Bächinger, *J. Biol. Chem.*, 1987, **262**, 17144-17148.

118. G. Fischer, B. Wittmann-Liebold, K. Lang, T. Kiefhaber and F. X. Schmid, *Nature*, 1989, **337**, 476-478.

119. L. A. Compton, J. M. Davies, J. R. Macdonald and H. P. Bächinger, *Eur. J. Biochem.*, 1992, **206**, 927-934.

120. P. L. Koser, G. P. Livi, M. A. Levy, M. Rosenberg and D. J. Bergsma, *Gene*, 1990, **96**, 189-195.

121. C. Hemenway and J. Heitman, *Ann. New York Acad. Sci.*, 1993, **696**, 38-46.

122. F. X. Schmid, L. M. Mayr, M. Mücke and E. R. Schönbrunner, *Adv. Protein Chem.*, 1993, **44**, 25-66.

123. M. Mücke and F. X. Schmid, *Biochemistry*, 1992, **31**, 7848-7854.

124. J. Fryman, E. Nimmesgern, K. Ohtsuka and F. U. Hartl, *Nature*, 1994, **370**, 111-117.

125. J. J. Siekierka, S. H. Y. Hung, M. Poe, C. S. Lin and N. H. Sigal, *Nature*, 1989, **341**, 755-757.

126. M. W. Harding, A. Galat, D. E. Uehling and S. L. Schreiber, *Nature*, 1989, **341**, 756-760.

127. R. K. Harrison and R. L. Stein, *Biochemistry*, 1990, **29**, 1684-1689.

128. R. K. Harrison and R. L. Stein, *Biochemistry*, 1990, **29**, 3813-3816.

129. S. L. Schreiber, *Science*, 1991, **251**, 283-287.

130. J. Liu and C. T. Walsh, *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 4028-4032.

131. S. L. Schreiber and D. B. Smith, *J. Org. Chem.*, 1989, **54**, 9-10.

132. M. Egbertson and S. J. Danishefsky, *J. Org. Chem.*, 1989, **54**, 11-12.

133. A. B. Jones, M. Yamaguchi, A. Patten, S. J. Danishefsky, J. A. Ragan, D. B. Smith and S. L. Schreiber, *J. Org. Chem.*, 1989, **54**, 17-19.

134. M. Nakatsuka, J. A. Ragan, T. Sammakia, D. B. Smith, D. E. Uehling and S. L. Schreiber, *J. Am. Chem. Soc.*, 1990, **112**, 5583-5601.

135. J. I. Luengo, A. L. Konialian and D. A. Holt, *Tetrahedron Lett.*, 1993, **34**, 991-994.

136. M. Furber, M. E. Cooper and D. K. Donald, *Tetrahedron Lett.*, 1993, **34**, 1351-1354.

137. M. B. M. Teunissen, F. H. M. Pistoor, H. A. H. Rongen, M. L. Kapsenberg and J. D. Bos, *Transplantation,* 1992, **53**, 875-881.

138. G. Fischer, *Angew. Chem. Int. Ed. Engl.,* 1994, **33**, 1415-1436.

139. D. Yang, M. K. Rosen and S. L. Schreiber, *J. Am. Chem. Soc.*, 1993, **115**, 819-820.

140. M. K. Rosen, D. Yang, P. K. Martin and S. L. Schreiber, *J. Am. Chem. Soc.*, 1993, **115**, 821-822.

141. H. Kessler, R. Haessner and W. Schüler, *Helv. Chim. Acta*, 1993, **76**, 117-130.

142. J. W. Becker, J. Rotonda, B. M. McKeever, H. K. Chan, A. I. Marcy, G. Wiederrecht, J. D. Hermes and J. P. Springer, *J. Biol. Chem.*, 1993, **268**, 11335-11339.

143. H. Ke, D. Mayrose, P. J. Belshaw, D. G. Alberg, S. L. Schreiber, Z. Y. Chang, F. A. Etzkorn, S. Ho and C. T. Walsh, *Structure,* 1994, **2**, 33-44.

144. R. P. Meadows, D. G. Nettesheim, R. X. Xu, E. T. Olejniczak, R. P. Petros, T. F. Holzman, J. Severin, E. Gubbins, H. Smith and S. W. Fesik, *Biochemistry*, 1993, **32**, 754-765.

145. L. T. Kakalis and I. M. Armitage, *Biochemistry,* 1994, **33**, 1495-1501.

146. R. L. Stein, *Adv. Protein Chem.,* 1993, **44**, 1-24.

147. G. Fischer, E. Berger and H. Bang, *FEBS Lett.,* 1989, **250**, 267-270.

148. J. Liu, M. W. Albers, C. -M. Chen, S. L. Schreiber and C. T. Walsh, *Proc. Natl. Acad. Sci. USA,* 1990, **87**, 2304-2308.

149. M. W. Albers, C. T. Walsh and S. L. Schreiber, *J. Org. Chem.*, 1990, **55**, 4984-4986.

150. R. K. Harrison and R. L. Stein, *J. Am. Chem. Soc.*, 1992, **114**, 3464-3471.

151. A. Giannis and T. Kolter, *Angew. Chem. Int. Ed. Engl.,* 1993, **32**, 1244-1267.

152. C. Barré, P. Le Grel, A. Robert and M. Baudy-Floc'h, *J. Chem. Soc., Chem. Comm.*, 1994, 607.

153. G. L. Olson, M. E. Voss, D. E. Hill, M. Kahn, V. S. Madison and C. M. Cook, *J. Am. Chem. Soc.,* 1990, **112**, 323-333.

154. M. Kahn, S. Wilke, B. Chen and K. Fujita, *J. Am. Chem. Soc.,* 1988, **110**, 1638-1639.

155. H. J. Dyson, M. Rance, R. A. Houghten, R. A. Lerner and P. E. Wright, *J. Mol. Biol.,* 1988, **201**, 161-200.

156. I. Z. Siemion, M. Lisowski, D. Konopinska and E. Nawrocka, *Eur. J. Biochem.,* 1980, **112**, 339-343.

157. M. G. Hinds, N. G. J. Richards and J. A. Robinson, *J. Chem. Soc., Chem. Comm.*, 1988, 1447-1449.

158. M. G. Hinds, J. H. Welsh, D. M. Brennand, J. Fisher, M. J. Glennie, N. G. J. Richards, D. L. Turner and J. A. Robinson, *J. Med. Chem.*, 1991, **34**, 1777-1789.

159. D. Gramberg and J. A. Robinson, *Tetrahedron Lett.*, 1994, **35**, 861-864.

160. T. P. Curran and P. M. McEnaney, *Tetrahedron Lett.*, 1995, **36**, 191-194.

161. P. K. C. Paul, P. A. Burney, M. M. Campbell and D. J. Osguthorpe, *Bioorg. Med. Chem. Lett.*, 1992, **2**, 141-144.

162. P. Juvvadi, D. J. Dooley, C. C. Humblet, G. H. Lu, E. A. Lunney, R. L. Panek, R. Skeean and G. R. Marshall, *Int. J. Peptide Protein Res.*, 1992, **40**, 163-170.

163. V. W. Magaard, R. M. Sanchez, J. W. Bean and M. L. Moore, *Tetrahedron Lett.*, 1993, **34**, 381-384.

164. M. Marraud, V. Dupont, V. Grand, S. Zerkout, A. Lecoq, G. Boussard, J. Vidal, A. Collet and A. Aubry, *Biopolymers*, 1993, **33**, 1135-1148.

165. A. Nefzi, K. Schenk and M. Mutter, *Prot. Pept. Lett.*, 1994, **1**, 66-69.

166. A. A. Abell, D. A. Hoult and E. J. Jamieson, *Tetrahedron Lett.*, 1992, **31**, 5831-5832.

167. G. D. Smith, J. Zabrocki, T. A. Flak and G. R. Marshall, *Int. J. Peptide Protein Res.*, 1991, **37**, 191-197.

168. L. W. Boteju, T. Zalewska, H. I. Yamamura and V. J. Hruby, *Bioorg. Med. Chem. Lett.*, 1993, **3**, 2011-2016.

169. S. F. Brady, W. J. Paleveda, B. H. Arison, R. Saperstein, E. J. Brady, K. Raynor, T. Reisine, D. F. Veber and R. M. Freidinger, *Tetrahedron*, 1993, **49**, 3449-3466.

170. C. J. Andres, T. L. Macdonald, T. D. Ocain and D. Langhi, *J. Org. Chem.*, 1993, **58**, 6609-6613.

171. C. Di Bello, A. Scatturin, G. D'Auria, M. Gargiulo, L. Paolillo, E. Trivellone and R. De Castiglione, *Biopolymers*, 1991, **31**, 643-652.

172. J. R. Cann, R. E. London, C. J. Unkefer, R. J. Vavrek and J. M. Stewart, *Int. J. Peptide Protein Res.*, 1987, **29**, 486-496.

173. E. L. Eliel and S. H. Wilen, *Stereochemistry of Organic Compounds*, John Wiley & Sons, New York, 1994, ch. 13.

174. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York, 1986.

175. H. Kessler, M. Klein and K. Wagner, *Int. J. Peptide Protein Res.*, 1988, **31**, 481-498.

176. R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149-2154.

177. G. Barany, N. Kneib-Cordonier and D. G. Mullen, *Int. J. Peptide Protein Res.*, 1987, **30**, 705-739.

178. R. B. Merrifield, *Makromol. Chem. Makromol. Symp.*, 1988, **19**, 31-67.

179. G. B. Fields and R. L. Noble, *Int. J. Peptide Protein Res.*, 1990, **35**, 161-214.

180. T. Akizawa, T. Ayabe, M. Matsukawa, M. Itoh, M. Nishi, H. Sato, M. Seiki and M. Yoshioka, *J. Prot. Chem.*, 1994, **13**, 483-484.

181. C. McInnes, C. M. Kay, R. S. Hodges and B. D. Sykes, *Biopolymers*, 1994, **34**, 1221-1230.

182. M. Bodansky and A. Bodansky, in *The Practice of Peptide Synthesis*, eds. K. Hafner, C. W. Rees, B. M. Trost, J. -M. Lehn, P. von Rague Schleyer and R. Zahradnik, Springer-Verlag, Berlin, 1984.

183. G. W. Anderson, J. E. Zimmerman and F. M. Callahan, *J. Am. Chem. Soc.*, 1967, **89**, 5012-5017.

184. F. M. F. Chen, Y. Lee, R. Steinauer and N. L. Benoiton, *Can. J. Chem.*, 1987, **65**, 613-618.

185. H. Waldmann and H. Kunz, *Liebigs. Ann. Chem.*, 1983, 1712-1725.

186. R. Deziel, *Tetrahedron Lett.*, 1987, **28**, 4371-4372.

187. T. -L. Ho, *Synth. Comm.*, 1978, **8**, 15-17.

188. A. Aubry, A. Abbadi, G. Boussard and M. Marraud, *New. J. Chem.*, 1987, **11**, 739-744.

189. G. T. Montelione, E. Arnold, Y. C. Meinwald, E. R. Stimson, J. B. Denton, S. -G. Huang, J. Clardy and H. A. Scheraga, *J. Am. Chem. Soc.*, 1984, **106**, 7946-7958.

190. Y. Wolman, A. Schejter and M. Sokolovsky, *J. Am. Chem. Soc.*, 1972, **94**, 1720-1723.

191. R. H. Boutin and G. M. Loudon, *J. Org. Chem.*, 1984, **49**, 4277-4284.

192. J. Coste, E. Frérot and P. Jouin, *Tetrahedron Lett.*, 1991, **32**, 1967-1970.

193. J. R. Spencer, V. V. Antonenko, N. G. J. Delaet and M. Goodman, *Int. J. Peptide Protein Res.*, 1992, **40**, 282-293.

194. G. T. Montelione, E. Arnold, Y. C. Meinwald, E. R. Stimson, J. B. Denton, S. -G. Huang, J. Clardy and H. A. Scheraga, *J. Am. Chem. Soc.*, 1984, **106**, 7946-7958.

195. R. Mueller and L. Revesz, *Tetrahedron Lett.*, 1994, **35**, 4091-4092.

196. C. Grathwohl and K. Wüthrich, *Biopolymers*, 1981, **20**, 2623-2633.

197. J. Jacobsen, W. Melander, G. Vaisnys and C. Horváth, *J. Phys. Chem.*, 1984, **88**, 4536-4542.

198. T. Drakenberg, K. -I. Dahlqvist and S. Forsén, *J. Phys. Chem.*, 1972, **76**, 2178-2183.

199. L. -N. Lin and J. F. Brandts, *Biochemistry*, 1983, 553-559.

200. R. E. Galardy and M. Liakopoulou-Kyriakides, *Int. J. Peptide Protein Res.*, 1982, **20**, 144-148.

201. G. Cignarella, F. Savelli, R. Cerri and P. Sanna, *J. Heterocycl. Chem.*, 1974, **11**, 1049-1053.

202. G. Cignarella, R. Cerri, F. Savelli, and A. Maselli, *J. Heterocycl. Chem.*, 1977, **14**, 465-468.

203. I. Z. Siemion, *Org. Mag. Res.*, 1971, **3**, 545-550.

204. I. Z. Siemion, *Liebigs. Ann. Chem.*, 1971, **748**, 88-95.

205. K. D. Kopple and M. Ohnishi, *J. Am. Chem. Soc.*, 1969, **91**, 962-970.

206. J. B. Hendrickson, *J. Am. Chem. Soc.*, 1961, **83**, 4537-4548.

207. K. Freudenberg and L. Markert, *Chem. Ber.*, 1927, **60**, 2447-2458.

208. A. Alexakis, N. Lensen and P. Mangeney, *Synlett*, 1991, 625-626.

209. R. Mozingo, *Org. Synth., Coll. Vol. III*, 1971, 181-183.

210. D. S. Kemp, M. D. Sidell and T. J. Shortridge, *J. Org. Chem.*, 1979, **44**, 4473-4476.

211. J. M. Mellor and N. M. Smith, *J. Chem. Soc., Perkin Trans. 1*, 1984, 2927-2931.

212. S. E. Denmark, O. Nicaise and J. P. Edwards, *J. Org. Chem.*, 1990, **55**, 6219-6223.

213. C. J. M. Stirling, *J. Chem. Soc.*, 1958, 4531-4536.

214. B. J. R. Nicolaus, E. Ballasio, G. Pagani, L. Mariani and E. Testa, *Helv. Chim. Acta*, 1965, **48**, 1867-1885.

215. M. J. Begley, L. Crombie, D. Haigh, R. C. F. Jones, S. Osborne and R. A. B. Webster, *J. Chem. Soc., Perkin Trans. 1*, 1993, 2027-2046.

216. U. Kramer, A. Guggisberg, M. Hesse and H. Schmid, *Angew. Chem. Int. Ed. Engl.*, 1977, **16**, 861-862.

217. H. Stach and M. Hesse, *Tetrahedron*, 1988, **44**, 1573-1590.

218. B. C. Barrass and D. T. Elmore, *J. Chem. Soc.*, 1957, 4830-4834.

219. H. T. Huang and C. Niemann, *J. Am. Chem. Soc.*, 1950, **72**, 921-922.

220. A. B. Mauger, R. B. Desai, I. Rittner and W. J. Rzeszotarski, *J. Chem. Soc., Perkin Trans. 1*, 1972, 2146-2148.

221. R. L. Hinman and D. Fulton, *J. Am. Chem. Soc.*, 1958, **80**, 1895-1900.

222. I. Tanaka, T. Kozima, T. Ashida, N. Tanaka and M. Kakudo, *Acta Cryst. B*, 1977, **33**, 116-119.

223. R. Balasubramanian, A. V. Lakshminarayanan, M. N. Sabesan, G. Tegoni, K. Venkatesan and G. N. Ramachandran, *Int. J. Protein Res.*,1971, **3**, 25-33.

224. H. Dorn, A. Zubek and K. Walter, *Liebigs Ann. Chem.*, 1967, **707**, 100-106.

225. F. E. Condon, *J. Org. Chem.*, 1972, **37**, 3615-3617.

226. N. P. Peet amd S. Sunder, *J. Org. Chem.*, 1975, **40**, 1909-1914.

227. N. P. Peet, S. Sunder and R. J. Cregge, *J. Org. Chem.*, 1976, **41**, 2733-2736.

228. C. T. Pedersen, *Acta Chem. Scand.*, 1964, **18**, 2199-2200.

229. R. J. De Feo and P. D. Strickler, *J. Am. Chem. Soc.*, 1963, **85**, 2915-2917.

230. E. Ronwin, *J. Am. Chem. Soc.*, 1952, **74**, 127-132.

231. S. -C. J. Fu, S. M. Birnbaum and J. P. Greenstein, *J. Am. Chem. Soc.*, 1954, **76**, 6054-6058.

232. B. F. Tawil, A. Guggisberg and M. Hesse, *Tetrahedron*, 1992, **48**, 3775-3780.

233. A. Lewis, *1st Year Transfer Report*, St. Andrews University, 1995.

234. R. D. Stipanovic and R. B. Turner, *J. Org. Chem.*, 1968, **33**, 3261-3263.

235. M. E. Jung, W. A. Andrus and P. L. Ornstein, *Tetrahedron Lett.*, 1977, 4175-4178.

236. E. E. Smissman, M. D. Corbett, N. A. Jenny and O. Kristiansen, *J. Org. Chem.*, 1972, **37**, 1700-1703.

237. M. P. Bueno, C. A. Cativiela and J. A. Mayoral, *J. Org. Chem.*, 1991, **56**, 6551-6555.

238. P. Balaram, *Curr. Opin. Struct. Biol.*, 1992, **2**, 845-851.

239. St. Guttmann, *Helv. Chim. Acta,* 1961, **44**, 721-744.

240. J. Vicar, J. Smolíková and K. Bláha, *Coll. Czech. Chem. Comm.,* 1972, **37**, 4060-4071.